

STIC-Biotech/ChemLib

78067

From: Rao, Manjunath N.
Sent: Thursday, October 17, 2002 12:57 PM
To: STIC-Biotech/ChemLib
Subject: Sequence search request for 09/687860

From: Manjunath N. Rao
Art Unit 1652, Room 10A11
Mail Box in Room ~~10C-01~~
Phone: 306-5681

→ 100001

CRFF

Date: 10-17-02

Please search the following as soon as possible for application with serial number **09/687,860**

SEQ ID NO:24, against all commercial protein databases including issued patents database and pending application database and provide a print of all results.

If y u have any questions please call me at the above phone number.

Thanks

Manjunath N. Rao, Ph.D.
Biotechnology Patent Examiner
Art Unit 1652, Room 10A11
Crystal Mall 1, USPTO.

Point of Contact:
Beverly Shears
Technical Info. Specialist
CM1 1E05 Tel: 308-4994

THIS PAGE BLANK (USPTO)

SEARCH REQUEST FORM

Requestor's

Name: _____

Serial

Number: _____

Date: _____

Phone: _____

Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

BEST AVAILABLE COPY

STAFF USE ONLY

Date completed: 10-18-02

Searcher: Beverly E 4994

Terminal time: 20

Elapsed time: _____

CPU time: _____

Total time: 23

Number of Searches: _____

Number of Databases: 1

Search Site

_____ STIC

_____ CM-1

_____ Pre-S

Type of Search

_____ N.A. Sequence

_____ A.A. Sequence

_____ Structure

_____ Bibliographic

Vendors

_____ IG Suite

_____ STN

_____ Dialog

_____ APS

_____ Geninfo

_____ SDC

_____ DARC/Questel

✓ Other CGN

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2002, 15:52:13 ; Search time 35 seconds
(without alignments)
1754.966 Million cell updates/sec

Title: US-09-687-860-24
Perfect score: 3060
Sequence: 1 MRLPMSIALPLLSTWVGAF.....VDGLVLTSGCLPDSLSTVDD 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapect 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3060	100.0	553	21	AAAB01423 Human TANGO 212.
2	3055	99.8	553	22	AAAG93622 Human polypeptide,
3	3049.5	99.7	554	22	AAAG93156 Human polypeptide,
4	3049.5	99.7	554	22	AAAB72224 Human EXMAD-2 SEQ
5	3030	99.0	553	20	AAV18108 Protein encoded by
6	2986.5	97.6	573	22	AAAG0942 Human polypeptide
7	2767	90.4	537	20	AAV18110 Protein encoded by
8	1931	63.1	338	20	AAV41702 Human PRO320 prote
9	1931	63.1	338	21	AAAB4258 Human PRO320 (UM2
10	1931	63.1	338	21	AAAB1869 Amino acid sequenc
11	1931	63.1	338	21	AAV95339 Human PRO320 antit

12	1894	61.9	331	22	AAAG23677
13	1491	48.7	333	21	AAAB01429
14	1288	42.1	284	21	AAV76110 Rat TGF-beta homol
15	1288	42.1	284	22	AAAB56049 Skin cell protein,
16	1105	36.1	546	22	AAAB70547 Clone 1646/945.0.8
17	1091.5	35.7	582	22	AAAB70547 protei
18	1030	33.7	509	20	AAV13397 Amino acid sequenc
19	1030	33.7	509	22	AAV29049 Human PRO polypept
20	1030	33.7	509	22	AAAB80265 Human PRO334 prote
21	976	31.9	251	22	ABBI0214 Human CDNA SEQ ID
22	976	31.9	251	22	AAU16935 Human novel secret
23	683	22.3	183	21	AAAB42085 Human ORFX ORF1849
24	661	21.6	211	22	AAAG04843 Human polypeptide
25	602	19.7	100	20	AAV18109 EGF motif containi
26	409.5	13.4	2912	22	ABG06402 Novel human diagno
27	393	12.8	1118	22	AAAG0209 Human fibrillin-11
28	372.5	12.2	2189	11	AAAR05222 Antigen GX5401FL e
29	369.5	12.1	333	21	AAV84709 Amino acid sequenc
30	367.5	12.0	636	22	AAAG0146 Human polypeptide
31	362.5	11.8	576	22	AAAB0174 Human protein SEQ
32	362.5	11.8	576	22	AAAB0175 Human protein SEQ
33	362.5	11.8	576	22	AAAG1932 Human polypeptide
34	362.5	11.8	603	22	AAAG1931 Human protein SEQ
35	362.5	11.8	650	22	AAAG79190 Human fibrillin 3.
36	357	11.7	1121	22	AAAG78887 Mouse latent TGF-b
37	354.5	11.6	1833	16	AAAR9478 Human fibrillin-11
38	354.5	11.6	1833	21	AAAB12270 Human fibrillin-11
39	350.5	11.5	497	22	AAAG50207 Human fibrillin type
40	350	11.4	683	18	AAAR1150 Human fibrillin type
41	350	11.4	683	18	AAAG27600 Novel human diagno
42	350	11.4	686	22	AAAG19385 Amino acid sequenc
43	348	11.4	585	22	AAAG7242 Human protein sequ
44	339.5	11.1	741	22	AAAG95002 Amino acid sequenc
45	339.5	11.1	780	22	AAAG7241 Amino acid sequenc

ALIGNMENTS

RESULT 1	AAAB01423 standard; Protein: 553 AA.
ID	AAAB01423
XX	AAAB01423:
AC	20-OCT-2000 (first entry)
XX	Human TANGO 212.
DE	TANGO: 128: 140: 197: 212: 213: 224: 239: modulating agent; asthma;
XX	graft versus-host diseases; rheumatoid arthritis; psoriasis;
KW	inflammatory bowel disease; septic shock; ulcerative colitis;
KW	Crohn's disease; chronic myelogenous leukemia; cancer; liver
KW	disease; Hodgkin's disease; osteoarthritis; Lyme's disease;
KW	cachexia; autoimmune disease; myasthenia gravis; autoimmune
KW	systemic lupus erythematosus; transgenic animal; diagnosis;
KW	prognosis; prophylactic; therapeutic; human.
XX	
OS	Homo sapiens.
XX	
PN	W0200039284-A1.
XX	
PD	06-JUL-2000.
XX	
PF	23-DEC-1999; 99WO-US31025.
XX	
PR	30-DEC-1998; 98US-0223546.
XX	
PA	(MILL-) MILLENNIUM PHARM INC.
XX	
PI	Holtzman DA;
XX	
DR	WPI: 2000-465743/40.

DR N-PSDB; AAA47456.
 XX Novel nucleic acid sequences encoding TANCO-128, 140, 197, 212, 213,
 PT 224 and 239 polypeptides useful for the treatment of asthma, rheumatoid
 PT arthritis, psoriasis and autoimmune diseases
 XX
 PS Claim 8; Fig 5; 209pp; English.
 XX
 CC Nucleic acids encoding TANCO polypeptides are useful as modulating
 CC agents for regulating cellular processes like asthma, graft
 CC versus-host diseases, rheumatoid arthritis, psoriasis, inflammatory
 CC bowel disease, septic shock, ulcerative colitis, Crohn's disease,
 CC chronic myelogenous leukemia, cancer, liver disease, Hodgkin's
 CC disease, osteoarthritis, Lyme's disease, cachexia and autoimmune
 CC diseases e.g. myasthenia gravis, autoimmune diabetes and systemic
 CC lupus erythematosus. The nucleic acids are also useful for producing
 CC transgenic animals and the TANCO polypeptides themselves. Partial
 CC TANCO-128, 140, 197, 212, 213, 224, 239 sequences are useful in
 CC forensic biology, for diagnostic assays, prognostic assays,
 CC pharmacogenomics and for monitoring clinical trials. TANCO
 CC polypeptides are suitable for both prophylactic and therapeutic
 CC methods for treating a subject at risk of a disorder or having a
 CC disorder associated with aberrant TANCO expression. A wide range
 CC of cellular disorders can be treated.
 CC
 XX
 XX Sequence 553 AA:

Query Match 100.0%; Score 3060; DB 21; Length 553;
 Best Local Similarity 100.0%; Pred. No. 1.8e-217;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFLPMSIALPRLLSVAVAGFGFNAASARHHGLASARQGVCHYGTAKLCCYGMRRNSKV 60
 DB 1 MFLPMSIALPRLLSVAVAGFGFNAASARHHGLASARQGVCHYGTAKLCCYGMRRNSKV 60
 QY 61 CEATEPCCKGEGCVGPKKRCFPGYTGKTCSDQVNECGMRRPQOHNCVNTHGSKFC 120
 DB 61 CEATEPCCKGEGCVGPKKRCFPGYTGKTCSDQVNECGMRRPQOHNCVNTHGSKFC 120
 QY 121 LSGHMLMPDATCVNSRTCAMINCOYSCEDTEBGPQCLPSSGLRLAPNGRDLIDECAS 180
 DB 121 LSGHMLMPDATCVNSRTCAMINCOYSCEDTEBGPQCLPSSGLRLAPNGRDLIDECAS 180
 QY 181 GKVICPYNRRCVNTFGSYCKCHIGFELQYISGRDCTIDINECTYDSTCHHANCFTQ 240
 DB 181 GKVICPYNRRCVNTFGSYCKCHIGFELQYISGRDCTIDINECTYDSTCHHANCFTQ 240
 QY 241 GSFCKCKOGYKGNLGRCSAIPENSVEVLRAPGTIKRIKLLAHKSMKKAKIKVT 300
 DB 241 GSFCKCKOGYKGNLGRCSAIPENSVEVLRAPGTIKRIKLLAHKSMKKAKIKVT 300
 QY 301 PEPTPTPPKYNLOPFNTEIYSRGNSHGKKGNEKKMEGLDEKREERALKNDIEER 360
 DB 301 PEPTPTPPKYNLOPFNTEIYSRGNSHGKKGNEKKMEGLDEKREERALKNDIEER 360
 QY 361 SLRGVFFPKVNEGEFLLIVORKALSKLEHKDNLNTSVCSFHNGLICDMKQDREDFD 420
 DB 361 SLRGVFFPKVNEGEFLLIVORKALSKLEHKDNLNTSVCSFHNGLICDMKQDREDFD 420
 QY 421 WNPADRDNAIGFYMAVPLAGHKDGRKLLPDLQPSNCLLFDYRLAGDKGKRLV 480
 DB 421 WNPADRDNAIGFYMAVPLAGHKDGRKLLPDLQPSNCLLFDYRLAGDKGKRLV 480
 QY 481 FVKNSSNNAIAEKTTSSEDEKMTGKIQLYOGTDATKSIIFEAERKGTGELAVDGLV 540
 DB 481 FVKNSSNNAIAEKTTSSEDEKMTGKIQLYOGTDATKSIIFEAERKGTGELAVDGLV 540
 QY 541 SGICPDSLVSVD 553
 DB 541 SGICPDSLVSVD 553

RESULT 2

AA093622
 ID AA093622 standard; Protein; 553 AA.
 XX
 AC AA093622;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide, SEQ ID NO: 3456.
 XX
 KW Human; full length cDNA; cDNA synthesis; oligo-capping.
 XX
 OS Homo sapiens.
 XX
 PN EP130094-A2.
 XX
 PD 05-SEP-2001.
 XX
 PE 07-JUL-2000; 2000EP-0114089.
 XX
 PR 08-JUL-1999; 99JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI: 2001-524255/58.
 XX N-PSDB; AAK94555.

830 Primers useful for synthesizing full length cDNA clones and their
 use in genetic manipulation -

Claim 8; SEQ ID NO 3456; 1380bp + sequence listing; English.

The invention relates to primers for synthesizing full length cDNA
 clones. 830 cDNA molecules encoding a human protein have been
 isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 molecules have been determined. Primers for synthesizing the full length
 cDNA are useful for clarifying the function of the protein encoded by
 the cDNA. The full length clones were obtained by construction of full
 length enriched cDNA libraries that were synthesised by the oligo-capping
 method. The primers enable the production of the full length cDNA easily
 without any special methods. The present sequence is a polypeptide
 encoded by a full length human cDNA of the invention.
 Note: The sequence data for this patent did not form part of the printed
 specification, but was obtained in CD-ROM format directly from EPO.

Sequence 553 AA:

Query Match 99.8%; Score 3055; DB 22; Length 553;
 Best Local Similarity 99.8%; Pred. No. 4.2e-217;
 Matches 552; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFLPMSIALPRLLSVAVAGFGFNAASARHHGLASARQGVCHYGTAKLCCYGMRRNSKV 60
 DB 1 MFLPMSIALPRLLSVAVAGFGFNAASARHHGLASARQGVCHYGTAKLCCYGMRRNSKV 60
 QY 61 CEATEPCCKGEGCVGPKKRCFPGYTGKTCSDQVNECGMRRPQOHNCVNTHGSKFC 120
 DB 61 CEATEPCCKGEGCVGPKKRCFPGYTGKTCSDQVNECGMRRPQOHNCVNTHGSKFC 120
 QY 121 LSGHMLMPDATCVNSRTCAMINCOYSCEDTEBGPQCLPSSGLRLAPNGRDLIDECAS 180
 DB 121 LSGHMLMPDATCVNSRTCAMINCOYSCEDTEBGPQCLPSSGLRLAPNGRDLIDECAS 180
 QY 181 GKVICPYNRRCVNTFGSYCKCHIGFELQYISGRDCTIDINECTYDSTCHHANCFTQ 240
 DB 181 GKVICPYNRRCVNTFGSYCKCHIGFELQYISGRDCTIDINECTYDSTCHHANCFTQ 240
 QY 241 GSFCKCKOGYKGNLGRCSAIPENSVEVLRAPGTIKRIKLLAHKSMKKAKIKVT 300
 DB 241 GSFCKCKOGYKGNLGRCSAIPENSVEVLRAPGTIKRIKLLAHKSMKKAKIKVT 300

```

Db 241 GSEFKCKQGYKNGGLRCSAIPENSVEVLRAPTIKDIRIKLLAHKNSMKKAKIKNVY 300
QY 301 PEPTRTPTPKVNIQPFNYEEIVSRGNSHGKKGNEEMKEGLEDEKREKALKNDIEER 360
Db 301 PEPTRTPTPKVNIQPFNYEEIVSRGNSHGKKGNEEMKEGLEDEKREKALKNDIEER 360
QY 361 SLRGDVEFFPKVNEAGEGGLILVORKALTSKLEHKDLNISVDCSFNHCICMWDREDDFD 420
Db 361 SLRGDVEFFPKVNEAGEGGLILVORKALTSKLEHKDLNISVDCSFNHCICMWDREDDFD 420
QY 421 WNPADRDNAIGFYMAVPALAGHKDKDGRKLKLLPDLOPQSNFCLLPYRLAGDKVGLRV 480
Db 421 WNPADRDNAIGFYMAVPALAGHKDKDGRKLKLLPDLOPQSNFCLLPYRLAGDKVGLRV 480
QY 481 FVKNNSNNALAMEKTTSEDEKMKTKGIQLYOGTDATKSIIFEAERGKGTGEIADVGLLV 540
Db 481 FVKNNSNNALAMEKTTSEDEKMKTKGIQLYOGTDATKSIIFEAERGKGTGEIADVGLLV 540
QY 541 SGLCPDLSLVDD 553
Db 541 SGLCPDLSLVDD 553

```

```

RESULT 3
AAM39156
ID AAM39156 standard; Protein; 554 AA.
XX
AC AAM39156;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 2301.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
peripheral nervous system; neuropathy; central nervous system; CNS;
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic;
chemokine; thrombolytic; drug screening; arthritis; inflammation;
leukemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000MO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
DR N-PSDB; AA158312.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 4; SEQ ID NO 2301; 10078bp; English.
XX
CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,

```

```

CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilization of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 554 AA;

```

```

Query Match 99.7%; Score 3049.5; DB 22; Length 554;
Best Local Similarity 99.8%; Pred. No. 1.1e-216;
Matches 553; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MRLPMSIALPRLLSWVAGFENNAASARRHGLLASAROPGVCHYGTAKLACCGWRNRSKGV 60
Db 1 MRLPMSIALPRLLSWVAGFENNAASARRHGLLASAROPGVCHYGTAKLACCGWRNRSKGV 60
QY 61 CEATCEPGCKRGECVGPNCRCFPGYTGKTSQDVNECGMPPRCQHRVTHGSYKFC 120
Db 61 CEATCEPGCKRGECVGPNCRCFPGYTGKTSQDVNECGMPPRCQHRVTHGSYKFC 120
QY 121 LSGHMLMPDATCVNSRCAMINCOYSCEDTEEGPQCLCPSSGLRLAPNGRDCLDIDECAS 180
Db 121 LSGHMLMPDATCVNSRCAMINCOYSCEDTEEGPQCLCPSSGLRLAPNGRDCLDIDECAS 180
QY 181 GKVICPYNRRCVNFFGSYCCCHIGFELYISGRYDCIDINECTMDSHTCSHHANCFTQ 240
Db 181 GKVICPYNRRCVNFFGSYCCCHIGFELYISGRYDCIDINECTMDSHTCSHHANCFTQ 240
QY 241 GSEFKCKQGYKNGGLRCSAIPENSVEVLRAPTIKDIRIKLLAHKNSMKKAKIKNVY 300
Db 241 GSEFKCKQGYKNGGLRCSAIPENSVEVLRAPTIKDIRIKLLAHKNSMKKAKIKNVY 300
QY 301 PEPTRTPTPKVNIQPFNYEEIVSRGNSHGKKGNEEMKEGLEDEKREKALKNDIEER 360
Db 301 PEPTRTPTPKVNIQPFNYEEIVSRGNSHGKKGNEEMKEGLEDEKREKALKNDIEER 360
QY 361 SLRGDVEFFPKVNEAGEGGLILVORKALTSKLEHK-DNISVDCSFNHCICMWDREDDFD 419
Db 361 SLRGDVEFFPKVNEAGEGGLILVORKALTSKLEHKADLNISVDCSFNHCICMWDREDDFD 420
QY 420 WNPADRDNAIGFYMAVPALAGHKDKDGRKLKLLPDLOPQSNFCLLPYRLAGDKVGLRV 479
Db 420 WNPADRDNAIGFYMAVPALAGHKDKDGRKLKLLPDLOPQSNFCLLPYRLAGDKVGLRV 480
QY 480 VEVKNSNNALAMEKTTSEDEKMKTKGIQLYOGTDATKSIIFEAERGKGTGEIADVGLLV 539
Db 480 VEVKNSNNALAMEKTTSEDEKMKTKGIQLYOGTDATKSIIFEAERGKGTGEIADVGLLV 540
QY 540 VSGLCPSDLSLVDD 553
Db 540 VSGLCPSDLSLVDD 554

```

```

RESULT 4
AAB27224
ID AAB27224 standard; Protein; 554 AA.
XX
AC AAB27224;
XX
DT 27-MAR-2001 (first entry)
XX
DE Human EXMAD-2 SEQ ID NO: 2.
XX
KW Extracellular matrix and adhesion-associated protein; EXMAD; cancer;

```

KM Inflammation; reproductive disorder; cardiovascular disorder;
 KM Immune disorder; musculoskeletal disorder; developmental disorder;
 KM Gastrointestinal disorder; cell proliferation disorder.
 OS Homo sapiens.
 XX WO200068380-A2.
 PN 16-NOV-2000.
 PD 10-MAY-2000; 2000WO-US12811.
 PF 11-MAY-1999; 99US-0133643.
 PR 23-AUG-1999; 99US-0150409.
 XX (INCY-) INCYTE GENOMICS INC.
 PA Bandman O, Hillman JL, Tang YT, Lai P, Yue H, Baughn MR, Lu DM;
 PI Azimzai Y;
 PI WPI: 2001-007395/01.
 DR N-PSDB; AAC66891.
 XX Isolated polynucleotide encoding extracellular matrix or
 PT adhesion-associated protein (EXMAD) useful for diagnosing, treating, or
 PT preventing disorders associated with expression of EXMAD such as
 PT proliferative, immune and genetic disorders -
 PS Claim 1; Page 88-89; 129pp; English.
 XX The present invention provides the protein and coding sequences for 25
 CC novel extracellular matrix and adhesion-associated proteins (EXMADs).
 CC These are designated EXMAD-1, EXMAD-2, EXMAD-3, EXMAD-4, EXMAD-5,
 CC EXMAD-6, EXMAD-7, EXMAD-8, EXMAD-9, EXMAD-10, EXMAD-11, EXMAD-12,
 CC EXMAD-13, EXMAD-14, EXMAD-15, EXMAD-16, EXMAD-17, EXMAD-18, EXMAD-19,
 CC EXMAD-20, EXMAD-21, EXMAD-22, EXMAD-23, EXMAD-24 and EXMAD-25. They are
 CC useful in the prevention and treatment of cancers, cell proliferation,
 CC cardiovascular, reproductive, immune, musculoskeletal, developmental and
 CC gastrointestinal disorders and inflammation.
 CC
 XX Sequence 554 AA;
 SO
 Query Match 99.7%; Score 3049.5; DB 22; Length 554;
 Best Local Similarity 99.8%; Pred. No. 1.1e-216;
 Matches 553; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 MPLPMSIALPLLISWVAGFGNMAASARHHGLASAROPGVCHYGTKLACCYGMRNSKGV 60
 DB 1 MPLPMSIALPLLISWVAGFGNMAASARHHGLASAROPGVCHYGTKLACCYGMRNSKGV 60
 QY 61 CEATCEPCRCRFGECVGNPKRCRCPFGYTGKTCSDVNECGMKPRPOHRCVTHGSKFC 120
 DB 61 CEATCEPCRCRFGECVGNPKRCRCPFGYTGKTCSDVNECGMKPRPOHRCVTHGSKFC 120
 QY 121 LSGHMLPADATCVNSRRCAMINCOYSCEDTREGQCICPSSGRLAPRGRCCLDIDECAS 180
 DB 121 LSGHMLPADATCVNSRRCAMINCOYSCEDTREGQCICPSSGRLAPRGRCCLDIDECAS 180
 QY 181 GKVICPNRRRCVNTFGSYGCHIGFELYISGRYDCIDINECTMDSTGSHNACFTQ 240
 DB 181 GKVICPNRRRCVNTFGSYGCHIGFELYISGRYDCIDINECTMDSTGSHNACFTQ 240
 QY 241 GSFCKCKOGYKGNLRCASAIPENSVKEVLAPGTIKDIRIKLLAHKNSMKKAKIKKAVT 300
 DB 241 GSFCKCKOGYKGNLRCASAIPENSVKEVLAPGTIKDIRIKLLAHKNSMKKAKIKKAVT 300
 QY 301 PEPTPTPPKYNVLOPENTEEIVSRGNSHGKKGNEEKMKEGLEDEKREERALKNDIEER 360
 DB 301 PEPTPTPPKYNVLOPENTEEIVSRGNSHGKKGNEEKMKEGLEDEKREERALKNDIEER 360
 QY 361 SLRGDVEFPKYNAGEGGLIYORKALTSKLEHK-DLNTISVDCSPFNHICOMKODREDF 419
 DB 361 SLRGDVEFPKYNAGEGGLIYORKALTSKLEHKADLNTISVDCSPFNHICOMKODREDF 420

QY 420 DWNPADRDNAIGFYMAVPALAGHKKIDIRKLIPDLPOSNFCLLPYRLAGDKVGLR 479
 DB 421 DWNPADRDNAIGFYMAVPALAGHKKIDIRKLIPDLPOSNFCLLPYRLAGDKVGLR 480
 QY 480 VEVKNSNNALAMEKTTSEDEKWKTKGIQLYOGDTATKSIIEAERKKGKGEIADVGLL 539
 DB 481 VEVKNSNNALAMEKTTSEDEKWKTKGIQLYOGDTATKSIIEAERKKGKGEIADVGLL 540
 QY 540 VSGLCPSDLSVDD 553
 DB 541 VSGLCPSDLSVDD 554
 -RESULT 5
 ID AAY18108 standard; Protein: 553 AA.
 AC AAY18108;
 AC AAY18108;
 DE 10-AUG-1999 (first entry)
 DE Protein encoded by cDNA insert of clone pEGFR-HY2.
 KW Epidermal growth factor; EGF repeat domain; haematopoiesis regulator;
 KW tissue growth activity; activin; inhibin; chemotaxis; chemokinesis;
 KW haemostasis; thrombolysis; anti-inflammatory; leukaemia; anaemia;
 KW immune disorder; immune deficiency; nervous system disorder; therapy.
 XX Synthetic.
 OS
 XX WO9927096-A1.
 PN 03-JUN-1999.
 PD 23-NOV-1998; 98WO-US24524.
 PP 22-NOV-1997; 97US-0968800.
 PR (HYSE-) HYSEQ INC.
 PA Crivenjakov R, Dickson M, Drmanac RT, Drmanac S;
 PI Ford J, Kita D, Labat I, Leshkowitz D;
 PI WPI: 1999-370904/31.
 DR N-PSDB; AAX79501.
 XX New polypeptide with epidermal growth factor repeat domains
 PS Claim 8; Fig 5; 96pp; English.
 CC This sequence represents a polypeptide of the invention, which has
 CC similarity to epidermal growth factor (EGF) repeat domains. The
 CC polypeptides and their compositions may have haematopoiesis regulating,
 CC tissue growth, activin/inhibin, chemotactic/chemokinetic, haemostatic,
 CC thrombolytic, receptor/ligand and anti-inflammatory activities. They may
 CC be used to treat leukemias, anaemias, immune disorders and deficiencies
 CC and nervous system disorders. They can be used in screening assays to
 CC identify agents which bind to them and the nucleotide sequences can be
 CC used as probes for in situ hybridisation. The polypeptides and their
 CC polynucleotides can also be used for other therapeutic, diagnostic and
 CC research utilities.
 CC
 XX Sequence 553 AA;
 SO
 Query Match 99.0%; Score 3030; DB 20; Length 553;
 Best Local Similarity 99.3%; Pred. No. 2.9e-215;
 Matches 549; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 MPLPMSIALPLLISWVAGFGNMAASARHHGLASAROPGVCHYGTKLACCYGMRNSKGV 60
 DB 1 MPLPMSIALPLLISWVAGFGNMAASARHHGLASAROPGVCHYGTKLACCYGMRNSKGV 60

Db 558 LTVSGLCPDLSLVD 573

RESULT 7
AA18110
ID AA18110 standard; Protein: 537 AA.
XX
XX AA18110;
XX
XX 10-AUG-1999 (first entry)
XX
XX Protein encoded by fragment of cDNA insert of clone pEGFR-HY2.
XX
XX Epidermal growth factor; EGF repeat domain; haematopoiesis regulator;
XX tissue growth activity; activin; inhibitor; chemotaxis; chemokinesis;
XX haemostasis; thrombolysis; anti-inflammatory; leukaemia; anaemia;
XX immune disorder; immune deficiency; nervous system disorder; therapy.
XX
XX Synthetic.
XX
XX WO9927096-A1.
XX
XX 03-JUN-1999.
XX
XX 23-NOV-1998; 98WO-US24524.
XX
XX 22-NOV-1997; 97US-0968800.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Crivenjakov R, Dickson M, Drmanac RT, Drmanac S;
XX Ford J, Kita D, Labat I, Leshkowitz D;
XX WPI: 1999-370904/31.
XX DR N-PSDB; AAX79503.
XX
XX New polypeptide with epidermal growth factor repeat domains
XX
XX Claim 21; Fig 3; 96pp; English.
XX
XX This sequence represents a polypeptide of the invention, which has
XX similarity to epidermal growth factor (EGF) repeat domains. The
XX polypeptides and their compositions may have haematopoiesis regulating,
XX tissue growth, activin/inhibin, chemotactic/chemokinetic, haemostatic,
XX thrombolytic, receptor/ligand and anti-inflammatory activities. They may
XX be used to treat leukaemias, anaemias, immune disorders and deficiencies
XX and nervous system disorders. They can be used in screening assays to
XX identify agents which bind to them and the nucleotide sequences can be
XX used as probes for in situ hybridisation. The polypeptides and their
XX polynucleotides can also be used for other therapeutic, diagnostic and
XX research utilities.
XX
XX Sequence 537 AA;
XX
XX Query Match 90.4%; Score 2767; DB 20; Length 537;
XX Best Local Similarity 99.6%; Pred. No. 7.1e-196;
XX Matches 500; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 52 GWRNSKGVCEATCEPGCKFGECVGNKRCRPGYTGTGKTSQDVNCGMKRPPCOHRCVN 111
DB 1 GWRNSKGVCEATCEPGCKFGECVGNKRCRPGYTGTGKTSQDVNCGMKRPPCOHRCVN 60
QY 112 THGSYKCFCLSGHMLMPATCVNSRTCAMINCQYSCDETEBPGQCLPSSGRLAPNGRD 171
DB 61 THGSYKCFCLSGHMLMPATCVNSRTCAMINCQYSCDETEBPGQCLPSSGRLAPNGRD 120
QY 172 CLDIDECASGVAVICPYNRCVNTFESYCKCHIGFELYISGRVYCIDINETMNSHCS 231
DB 121 CLDIDECASGVAVICPYNRCVNTFESYCKCHIGFELYISGRVYCIDINETMNSHCS 180
QY 232 HHANCFNTQGSFKCKCKQGYKNGJLRCASAIPENSYKVELRAFGTITKDRITKLAKHNSMK 291
DB 181 HHANCFNTQGSFKCKCKQGYKNGJLRCASAIPENSYKVELRAFGTITKDRITKLAKHNSMK 240

QY 292 KKAQIKNVTPEPTPTPEKVINLOPFNYEIRVSRGNSHGKKGNEKMEKEGLEDKREK 351
DB 241 KKAQIKNVTPEPTPTPEKVINLOPFNYEIRVSRGNSHGKKGNEKMEKEGLEDKREK 300
QY 352 ALKNDIEERSLGDVFFPKVNEAGEFGLIVORKALTSKLEKDLNISVDCSFNHCIDW 411
DB 301 ALKNDIEERSLGDVFFPKVNEAGEFGLIVORKALTSKLEKDLNISVDCSFNHCIDW 360
QY 412 KODREDDPDMNPADBDNALGFMAVPALAGHKKDIGRLKLLPDLOPOSFCLLPDYRLA 471
DB 361 KODREDDPDMNPADBDNALGFMAVPALAGHKKDIGRLKLLPDLOPOSFCLLPDYRLA 420
QY 472 GDKVKRLRYFVKNNSNALAMEKTSDEKWKTKGIQLYGTDATKSIIFEARGKGTGE 531
DB 421 GDKVKRLRYFVKNNSNALAMEKTSDEKWKTKGIQLYGTDATKSIIFEARGKGTGE 480
QY 532 IAVDGYLVSGICPDSLSVD 553
DB 481 IAVDGYLVSGICPDSLSVD 502

RESULT 8
AA41702
ID AA41702 standard; Protein: 338 AA.
XX
XX AA41702;
XX
XX 07-DEC-1999 (first entry)
XX
XX Human PRO320 protein sequence.
XX
XX Human; PRO: EST: expressed sequence tag; PCR primer; hybridisation;
XX KW probe; blood coagulation disorder; cancer; cellular adhesion disorder;
XX secreted protein; transmembrane protein.
XX
XX Homo sapiens.
XX
XX OS
XX PN WO9946281-A2.
XX
XX 16-SEP-1999.
XX
XX 08-MAR-1999; 99WO-US05028.
XX
XX PF
XX 10-MAR-1998; 98US-0077450.
XX PR 11-MAR-1998; 98US-0077632.
XX PR 11-MAR-1998; 98US-0077641.
XX PR 11-MAR-1998; 98US-0077649.
XX PR 12-MAR-1998; 98US-0077791.
XX PR 13-MAR-1998; 98US-0078004.
XX PR 17-MAR-1998; 98US-0040220.
XX PR 20-MAR-1998; 98US-0078886.
XX PR 20-MAR-1998; 98US-0078910.
XX PR 20-MAR-1998; 98US-0078936.
XX PR 20-MAR-1998; 98US-0078939.
XX PR 25-MAR-1998; 98US-0079294.
XX PR 25-MAR-1998; 98US-0079656.
XX PR 26-MAR-1998; 98US-0079663.
XX PR 27-MAR-1998; 98US-0079664.
XX PR 27-MAR-1998; 98US-0079689.
XX PR 27-MAR-1998; 98US-0079728.
XX PR 27-MAR-1998; 98US-0079786.
XX PR 30-MAR-1998; 98US-0079920.
XX PR 30-MAR-1998; 98US-0079923.
XX PR 31-MAR-1998; 98US-0080105.
XX PR 31-MAR-1998; 98US-0080107.
XX PR 31-MAR-1998; 98US-0080165.
XX PR 31-MAR-1998; 98US-0080194.
XX PR 01-APR-1998; 98US-0080327.
XX PR 01-APR-1998; 98US-0080328.
XX PR 01-APR-1998; 98US-0080333.
XX PR 01-APR-1998; 98US-0080334.
XX PR 08-APR-1998; 98US-0081049.

PR 08-APR-1998; 9805-0081070.
 PR 08-APR-1998; 9805-0081071.
 PR 09-APR-1998; 9805-0081195.
 PR 09-APR-1998; 9805-0081203.
 PR 09-APR-1998; 9805-0081229.
 PR 15-APR-1998; 9805-0081817.
 PR 15-APR-1998; 9805-0081838.
 PR 15-APR-1998; 9805-0081952.
 PR 15-APR-1998; 9805-0081955.
 PR 21-APR-1998; 9805-0082568.
 PR 21-APR-1998; 9805-0082569.
 PR 22-APR-1998; 9805-0082700.
 PR 22-APR-1998; 9805-0082704.
 PR 23-APR-1998; 9805-0082804.
 PR 23-APR-1998; 9805-0082767.
 PR 27-APR-1998; 9805-0082756.
 PR 28-APR-1998; 9805-0083336.
 PR 28-APR-1998; 9805-0083342.
 PR 29-APR-1998; 9805-0083392.
 PR 29-APR-1998; 9805-0083495.
 PR 29-APR-1998; 9805-0083496.
 PR 29-APR-1998; 9805-0083499.
 PR 29-APR-1998; 9805-0083500.
 PR 29-APR-1998; 9805-0083545.
 PR 29-APR-1998; 9805-0083554.
 PR 29-APR-1998; 9805-0083558.
 PR 29-APR-1998; 9805-0083559.
 PR 30-APR-1998; 9805-0083742.
 PR 05-MAY-1998; 9805-0084366.
 PR 06-MAY-1998; 9805-0084414.
 PR 06-MAY-1998; 9805-0084441.
 PR 07-MAY-1998; 9805-0084598.
 PR 07-MAY-1998; 9805-0084600.
 PR 07-MAY-1998; 9805-0084627.
 PR 07-MAY-1998; 9805-0084637.
 PR 07-MAY-1998; 9805-0084639.
 PR 07-MAY-1998; 9805-0084640.
 PR 07-MAY-1998; 9805-0084643.
 PR 13-MAY-1998; 9805-0085323.
 PR 13-MAY-1998; 9805-0085338.
 PR 13-MAY-1998; 9805-0085339.
 PR 15-MAY-1998; 9805-0085573.
 PR 15-MAY-1998; 9805-0085579.
 PR 15-MAY-1998; 9805-0085580.
 PR 15-MAY-1998; 9805-0085582.
 PR 15-MAY-1998; 9805-0085689.
 PR 15-MAY-1998; 9805-0085697.
 PR 15-MAY-1998; 9805-0085700.
 PR 15-MAY-1998; 9805-0085704.
 PR 18-MAY-1998; 9805-0086023.
 PR 22-MAY-1998; 9805-0086392.
 PR 23-MAY-1998; 9805-0086414.
 PR 23-MAY-1998; 9805-0086430.
 PR 22-MAY-1998; 9805-0086486.
 PR 28-MAY-1998; 9805-0087098.
 PR 28-MAY-1998; 9805-0087106.
 PR 28-MAY-1998; 9805-0087208.
 PR 30-JUL-1998; 9805-0094651.
 PR 11-SEP-1998; 9805-0100038.
 XX (GETH) GENENTECH INC.
 PA
 PI Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
 XX WPI: 1999-551358/46.
 DR N-PSDB; AA233991.
 XX
 PT New secreted and transmembrane polypeptides and their polynucleotides,
 PT adhesion disorders -
 XX
 PS Claim 12; Fig 45; 53opp; English.
 XX

CC The present invention describes secreted and transmembrane polypeptides
 CC and their polynucleotides. The nucleotide sequences are useful as
 CC sources of probes, primers, for chromosome mapping, and for generation
 CC of antisense sequences. They can also be used to create transgenic
 CC animals. The proteins can be used to treat a variety of diseases and
 CC disorders, depending on their function. Diseases that may be treated
 CC include blood coagulation disorders, cancers and cellular adhesion
 CC disorders. They may also be used to raise antibodies. AA233891 to
 CC AA234338; and AA41685 to AA41774 represent polynucleotide and
 CC polypeptide sequence given in the exemplification of the present
 CC invention.
 CC
 XX Sequence 338 AA;
 XX
 SQ
 Query Match 63.1%; Score 1931; DB 20; Length 338;
 Best Local Similarity 100.0%; Pred. No. 1,9e-134;
 Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MFLPMSLALPRLLSVWAGFGMAASARRHHGLASAROPGVCHYGTKLACCYGMRNSKGV 60
 Db 1 MFLPMSLALPRLLSVWAGFGMAASARRHHGLASAROPGVCHYGTKLACCYGMRNSKGV 60
 QY 61 CEATCEPCKFGECVGNPKCRCPGYTGKTSQDVNECGMKPRPCQHRCVNTHGSYKFC 120
 Db 61 CEATCEPCKFGECVGNPKCRCPGYTGKTSQDVNECGMKPRPCQHRCVNTHGSYKFC 120
 QY 121 LSGHMLPMDATCVNSRCAMNCOYSCEDTEEGQCLCPSSGLRLAPGRCDLIDECAS 180
 Db 121 LSGHMLPMDATCVNSRCAMNCOYSCEDTEEGQCLCPSSGLRLAPGRCDLIDECAS 180
 QY 181 GKVICPYNRCVNTFGSYKCHIGFELQYISGRYCIDINECTMDSHSHANCFNTQ 240
 Db 181 GKVICPYNRCVNTFGSYKCHIGFELQYISGRYCIDINECTMDSHSHANCFNTQ 240
 QY 241 GSFCKCKQGYKNGLRCSAIPENSVEVLRAPGTIDRIKLLAHKNSMKKKAKIKNT 300
 Db 241 GSFCKCKQGYKNGLRCSAIPENSVEVLRAPGTIDRIKLLAHKNSMKKKAKIKNT 300
 QY 301 PEPTRTPPKYNLQPFNVEEIVSRGNSHGKKGNEK 338
 Db 301 PEPTRTPPKYNLQPFNVEEIVSRGNSHGKKGNEK 338
 RESULT 9
 AAB44258
 ID AAB44258 standard; Protein: 338 AA.
 XX
 AC AAB44258;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human PRO320 (UNQ281) protein sequence SEQ ID NO:119.
 XX
 KW Human; secreted protein; transmembrane protein; PRO; EST; cytosolic;
 XX expressed sequence tag; detection; cancer.
 OS Homo sapiens.
 XX
 PN W0200053756-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 18-FEB-2000; 2000WO-US04341.
 XX
 PR 08-MAR-1999; 99NO-US05028.
 PR 12-MAR-1999; 99US-0123957.
 PR 29-MAR-1999; 99US-0126773.
 PR 21-APR-1999; 99US-0130232.
 PR 28-APR-1999; 99US-0131445.
 PR 14-MAY-1999; 99US-0134287.
 PR 23-JUN-1999; 99US-0141037.
 PR 26-JUL-1999; 99US-0145698.
 PR 29-OCT-1999; 99US-0162506.

PR 30-NOV-1999; 99WO-US28313.
 PR 02-DEC-1999; 99WO-US28551.
 PR 02-DEC-1999; 99WO-US28565.
 PR 16-DEC-1999; 99WO-US30095.
 PR 30-DEC-1999; 99WO-US31243.
 PR 30-DEC-1999; 99WO-US31274.
 PR 05-JAN-2000; 2000WO-US00219.
 PR 06-JAN-2000; 2000WO-US00277.
 PR 06-JAN-2000; 2000WO-US00376.
 XX
 PA (GETH) GENENTECH INC.
 PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
 PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
 PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
 PI Kljavin LJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA;
 PI Shelton DL, Stewart TA, Tumas D, Williams PM, Wood WI;
 XX
 XX WPI; 2000-611443/58.
 DR N-PSDB; AAC78484.
 XX
 PT Novel PRO polypeptides and polynucleotides used in detection methods,
 PT to target bioactive molecules to specific cells, and to modulate
 PT cellular activities -
 PS
 PS Claim 12; Fig 45; 636pp; English.
 XX
 CC AAC78458 to AAC78599 represent polynucleotide and EST (expressed
 CC sequence tag) sequences which encode secreted or transmembrane PRO
 CC polypeptides. The PRO polynucleotides and polypeptides have cytosstatic
 CC activity. The polynucleotides and polypeptides can be used for detecting
 CC the presence of PRO polypeptides in samples, for linking bioactive
 CC molecules to cells and for modulating biological activities of cells,
 CC using the polypeptides for specific targeting. The polypeptide targeting
 CC can be used to kill the target cells, e.g. for the treatment of cancers.
 CC The polypeptide pairs provide specific targeting of bioactive molecules
 CC to cells. AAC78600 to AAC78987 represent PCR primers and probes used in
 CC the isolation of the PRO polynucleotide sequences.
 CC
 XX
 SQ Sequence 338 AA;
 Query Match 63.1%; Score 1931; DB 21; Length 338;
 Best Local Similarity 100.0%; Pred. No. 1.9e-134;
 Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRLPMSIALPRLLSVWAGFGNNAASARHHGLASAROPGVCHYGTGKLACCYGWRNSKGV 60
 DB 1 MRLPMSIALPRLLSVWAGFGNNAASARHHGLASAROPGVCHYGTGKLACCYGWRNSKGV 60
 QY 61 CEATCEPCCKGCEVGPYKRCRCPFGYTGKTSQDVNECGMKRPRCOHRCVNTHGSKCFC 120
 DB 61 CEATCEPCCKGCEVGPYKRCRCPFGYTGKTSQDVNECGMKRPRCOHRCVNTHGSKCFC 120
 QY 121 LSGHMLMPDAPCVNSRRTCAMINCOVSCDETEBGPQCLCPSSGLRLAPRGROLDIDECAS 180
 DB 121 LSGHMLMPDAPCVNSRRTCAMINCOVSCDETEBGPQCLCPSSGLRLAPRGROLDIDECAS 180
 QY 181 GKVICPYNRRCVNTFGSYKCHIGFELQYISGRYCDIDINECTMDSHSTGSHNACFNTO 240
 DB 181 GKVICPYNRRCVNTFGSYKCHIGFELQYISGRYCDIDINECTMDSHSTGSHNACFNTO 240
 QY 241 GSFCKCKQGYKGNLRCSAIPENSVKQVLAAPGTIKRIKLLAHKNSMKKAKIKVNT 300
 DB 241 GSFCKCKQGYKGNLRCSAIPENSVKQVLAAPGTIKRIKLLAHKNSMKKAKIKVNT 300
 QY 301 PEPTRTPTPKVNLDPFNVEEIVSRGNSHGKKGNEK 338
 DB 301 PEPTRTPTPKVNLDPFNVEEIVSRGNSHGKKGNEK 338

XX AAB18669;
 AC 22-JAN-2001 (first entry)
 XX
 DT
 XX Amino acid sequence of a human PRO320 polypeptide.
 DE
 XX Fibulin homologue; PRO320; PRO938; PRO1031; PRO296; PRO213; PRO1330;
 KW PRO1449; angiogenesis; cardiovascularisation; cardiovascular disorder;
 KW endothelial disorder; angiogenic disorder; cancer; trauma; wound;
 KW arteriosclerosis; cardiac hypertrophy.
 XX
 XX Homo sapiens.
 OS
 XX
 XX Key
 FH Peptide
 FT Location/Qualifiers
 FT 1..21
 FT /note= "signal sequence"
 FT 18..24
 FT /note= "N-myristoylation site"
 FT 21..27
 FT /note= "N-myristoylation site"
 FT 30..36
 FT /note= "N-myristoylation site"
 FT 44..50
 FT /note= "N-myristoylation site"
 FT 54..58
 FT /note= "cAMP- and cGMP-dependent protein kinase phosphorylation site"
 FT 59..65
 FT /note= "N-myristoylation site"
 FT 68..74
 FT /note= "N-myristoylation site"
 FT 80..92
 FT /note= "EGF-like domain cysteine pattern signatu"
 FT 109..121
 FT /note= "aspartic acid and asparagine hydroxylation site"
 FT 114..120
 FT /note= "N-myristoylation site"
 FT 191..203
 FT /note= "aspartic acid and asparagine hydroxylation site"
 FT 196..202
 FT /note= "N-myristoylation site"
 FT 236..248
 FT /note= "aspartic acid and asparagine hydroxylation site"
 FT 241..247
 FT /note= "N-myristoylation site"
 FT 255..261
 FT /note= "N-myristoylation site"
 FT 326..332
 FT /note= "N-myristoylation site"
 FT 330..336
 FT /note= "N-myristoylation site"
 FT 330..334
 FT /note= "amidation site"
 XX
 PN W0200053752-A2.
 XX
 PD 14-SEP-2000.
 PD
 XX 30-DEC-1999; 99WO-US31274.
 PE
 XX
 XX 08-MAR-1999; 99WO-US05028.
 PR 21-APR-1999; 99US-0130232.
 PR 26-APR-1999; 99US-0131022.
 PR 28-APR-1999; 99US-0131445.
 PR 14-MAY-1999; 99US-0134287.
 PR 02-DEC-1999; 99WO-US28565.
 XX
 XX (GETH) GENENTECH INC.
 PA Baker KP, Ferrara N, Goddard A, Gurney AL, Hillan KJ, Williams PM;
 PI Wood WI;
 XX

DR WPI: 2000-638138/61.
 DR N-PSDB; AAAY5686.
 XX
 PT A composition useful for treatment and diagnosis of a cardiovascular,
 PT endothelial or angiogenic disorder, especially cancer, comprises (an
 PT agonist or antagonist of) a PRO320, PRO938, PRO1031, PRO296, PRO213,
 PT PRO1330 or PRO1449 polypeptide -
 XX
 PS Claim 67; Fig 2; 152pp; English.
 XX
 CC The present sequence represents PRO320, a fibulin homologue. The
 CC specification describes PRO320, PRO938, PRO1031, PRO296, PRO213,
 CC PRO1330 and PRO1449 polypeptides. The polypeptides promoter or
 CC inhibit angiogenesis and cardiovascularisation in mammals. The
 CC polypeptides are used for the treatment and diagnosis of a
 CC cardiovascular, endothelial or angiogenic disorder, especially
 CC cancer. Disorders that can be diagnosed, treated or prevented by
 CC the polypeptides of the invention include trauma such as wounds,
 CC arteriosclerosis, and cardiac hypertrophy.
 CC
 XX
 SQ Sequence 338 AA;
 Query Match 63.1%; Score 1931; DB 21; Length 338;
 Best Local Similarity 100.0%; Pred. No. 1.9e-134;
 Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLPMSIALPLLSSWVAGFGNNAASARRHGLLASARQPGVCHYGTAKLACCYGMRRNSKGV 60
 DB 1 MPLPMSIALPLLSSWVAGFGNNAASARRHGLLASARQPGVCHYGTAKLACCYGMRRNSKGV 60
 QY 61 CEATCEGCKGFECEVCVGNKRCRPFYGTGKTCSQDVNECGMKRRCQHRCAVTHGSYKFC 120
 DB 61 CEATCEGCKGFECEVCVGNKRCRPFYGTGKTCSQDVNECGMKRRCQHRCAVTHGSYKFC 120
 QY 121 LSGHMLMPDATCVNSRRCAMINCOYSCEDTEEGPQCLCPSSGLRLAPNGRCLDIDECAS 180
 DB 121 LSGHMLMPDATCVNSRRCAMINCOYSCEDTEEGPQCLCPSSGLRLAPNGRCLDIDECAS 180
 QY 181 GKVCPCPNRRRCVNTFGSYCCCHIGFELYISGRYDCIDINECMDSHTCSHHANCFWQ 240
 DB 181 GKVCPCPNRRRCVNTFGSYCCCHIGFELYISGRYDCIDINECMDSHTCSHHANCFWQ 240
 QY 241 GSFCKCKGKGYKNGLCASALPENSVEVLRAPCTIKDRIKLLAHKNSMKKAKIKNV 300
 DB 241 GSFCKCKGKGYKNGLCASALPENSVEVLRAPCTIKDRIKLLAHKNSMKKAKIKNV 300
 QY 301 PEPTRTPTPKVNLQPFNVEEIVSRGNSHGKKGNEEK 338
 DB 301 PEPTRTPTPKVNLQPFNVEEIVSRGNSHGKKGNEEK 338

RESULT 11
 AAY95339
 ID AAY95339 standard; Protein; 338 AA.
 XX
 AC AAY95339;
 XX
 DT 25-SEP-2000 (first entry)
 XX
 DE Human PRO320 antitumour protein.
 XX
 KW PRO320; human; antitumour; tumour; therapy; cytostatic;
 KW breast cancer; ovarian cancer; renal cancer; colorectal cancer;
 KW uterine cancer; prostate cancer; lung cancer; bladder cancer;
 KW central nervous system cancer; melanoma; leukemia; neoplasm.
 XX
 OS Homo sapiens.
 XX
 FH Key location/Qualifiers
 FT Peptide 1..21
 FT Protein /label= Signal_peptide
 FT 22..338
 FT /label= PRO320

FT Modified-site 330..334
 FT /note= "amidation site"
 FT
 FT Modified-site 109..121
 FT /note= "Asn hydroxylation site"
 FT
 FT Modified-site 191..203
 FT /note= "Asn hydroxylation site"
 FT
 FT Modified-site 236..248
 FT /note= "Asn hydroxylation site"
 FT
 FT Region 80..91
 FT /note= "epidermal growth factor-like domain
 FT /note= "cysteine pattern signature"
 FT
 FT Domain 103..125
 FT /note= "calcium-binding epidermal growth
 FT /note= "factor-like domain"
 FT
 FT Domain 230..252
 FT /note= "calcium-binding epidermal growth
 FT /note= "factor-like domain"
 FT
 FT Domain 185..207
 FT /note= "calcium-binding epidermal growth
 FT /note= "factor-like domain"
 FT
 XX
 PN W0200037638-A2.
 XX
 PD 29-JUN-2000.
 XX
 PF 02-DEC-1999; 99WO-US28565.
 XX
 PR 22-DEC-1998; 98US-0113296.
 PR 08-MAR-1999; 99WO-US05028.
 PR 21-APR-1999; 99US-0130232.
 PR 28-APR-1999; 99US-0131445.
 PR 14-MAY-1999; 99US-0134287.
 PR 20-JUL-1999; 99US-0144758.
 PR 26-JUL-1999; 99US-0145638.
 PR 15-SEP-1999; 99WO-US21090.
 PR 15-SEP-1999; 99WO-US21547.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;
 PI Napier MA, Pitti RM, Wood WL;
 PI
 DR WPI: 2000-442668/38.
 DR N-PSDB; AAA49718.
 XX
 PT Novel composition to inhibit neoplastic cell growth or for treating
 PT tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,
 PT PRO221, PRO324, PRO328, PRO301, PRO356, PRO359 or
 PT PRO866 -
 XX
 PS Claim 19; Fig 6; 172pp; English.
 XX
 CC The present sequence is that of human antitumour protein PRO320, as
 CC deduced from a foetal lung cDNA clone (see AAA49718). PRO320 has a
 CC mol.wt. of 37,143 and a pI of 8.92. A claimed method for inhibiting
 CC the growth of a tumour cell comprises exposing the tumor cell
 CC to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224, PRO328, PRO301,
 CC PRO356, PRO359 or PRO866 (see AAY95337-49), their
 CC agonists or chimeric polypeptides incorporating them. The tumour
 CC is especially a cancer selected from breast, ovarian, renal,
 CC colorectal, uterine, prostate, lung, bladder and central nervous
 CC system cancer, melanoma and leukemia. Methods for the recombinant
 CC expression of the antitumour proteins are also provided.
 XX
 SQ Sequence 338 AA;
 Query Match 63.1%; Score 1931; DB 21; Length 338;
 Best Local Similarity 100.0%; Pred. No. 1.9e-134;
 Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLPMSIALPLLSSWVAGFGNNAASARRHGLLASARQPGVCHYGTAKLACCYGMRRNSKGV 60
 DB 1 MPLPMSIALPLLSSWVAGFGNNAASARRHGLLASARQPGVCHYGTAKLACCYGMRRNSKGV 60

Qy	61	CEATCEBGCCKGECVGPAPKCKCPFGYATGKTCSODVNCGMKPRPCOHRCAVNTGSKYKFC	120
Dp	61	CEATCEBCKGECVGPAPKCKCPFGYATGKTCSODVNECGMKPRPCOHRCAVNTGSKYKFC	120
Qy	121	LSGHMLMPDATCVNSRTCCAMINCQYSCDEPTEEGPOGLCPSSGRLTAPANGDCDIDIECAS	180
Dp	121	LSGHMLMPDATCVNSRTCCAMINCQYSCDEPTEEGPOGLCPSSGRLTAPANGDCDIDIECAS	180
Qy	181	GKVICPPNRRCVNFFGSYYCCCHTGFELQYTSGRYDCDIDINECMSSHSTSHANCFNTQ	240
Dp	181	GKVICPPNRRCVNFFGSYYCCCHTGFELQYTSGRYDCDIDINECMSSHSTSHANCFNTQ	240
Qy	241	GSFCKCKCKOGYKGNGLRCSAIPENSVKRYELRAPETIDRIKLTLAHRNSMKKRAKIKNTY	300
Dp	241	GSFCKCKCKOGYKGNGLRCSAIPENSVKRYELRAPETIDRIKLTLAHRNSMKKRAKIKNTY	300
Qy	301	PEPRTPTPPKYNLOPENTYEEYVNSGGNSHGCKGKNGNEK	338
Dp	301	PEPRTPTPPKYNLOPENTYEEYVNSGGNSHGCKGKNGNEK	338

RESULT 12	
AA23677	
ID	AA23677 standard; Protein; 331 AA.
XX	
AC	AA23677;
XX	
DT	12-OCT-2001 (first entry)
XX	
DE	Human EST encoded protein SEQ ID NO: 1202.
XX	
KW	Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW	tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW	diagnostics; forensic test; gene mapping; genetic disorder;
KW	biodiversity; gene therapy; nutrition.
XX	
OS	Homo sapiens.
XX	
MO	200154477-A2.
XX	
PD	02-AUG-2001.
XX	
PE	25-JAN-2001; 2001WO-US02687.
XX	
PR	25-JAN-2000; 2000US-0491404.
PR	17-JUL-2000; 2000US-0617746.
PR	03-AUG-2000; 2000US-0631451.
PR	15-SEP-2000; 2000US-0663870.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
PI	Cao Y, Drmanac RA, Zhang J, Werhman T;
DR	WPI; 2001-476164/51.
DR	N-PSDB; AAH98336.
XX	
PT	Isolated polypeptide for treatment of diseases, diagnostics, raising
PT	antibodies and research use -
XX	
PS	Claim 20; Page 868-869; 1275pp; English.
XX	
CC	The present invention provides the protein and coding sequences of novel
CC	proteins from a variety of organisms, including human, dog, cat, horse,
CC	cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC	urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC	from the organism of interest. They can be used in diagnostics,
CC	forensics, gene mapping, identification of mutations, to assess
CC	biodiversity and for nutritional purposes. The present sequence is a
CC	protein of the invention.
XX	
XX	Sequence 331 AA;

Query Match	61.9%;	Score 1894;	DB 22;	Length 331;
Best Local Similarity	100.0%;	Pred. No. 9.7e-132;		
Matches 331;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	MPHPSIALPILLSVMVAGFENMAASRHHGILLASAROPGVCHNGTKLACCYGMRRNSKV	60
Db	1	MPPLMSIALPLILLSVMVAGFENMAASARHHGILLASAROPGVCHNGTKLACCYGMRRNSKV	60
QY	61	CEATCEPCKFGECEVGPBKRCRFPGYTGKTCSDOVNECGMKPRPCOHRCVNTHGYSKFC	120
Db	61	CEATCEPCKFGECEVGPBKRCRFPGYTGKTCSDOVNECGMKPRPCOHRCVNTHGYSKFC	120
QY	121	LSGHMLMPDADCVSRRICAMINCOYSCEDPFEPSQCCLPSSGRLANNGDCLDIDBCAS	180
Db	121	LSGHMLMPDADCVSRRICAMINCOYSCEDPFEPSQCCLPSSGRLANNGDCLDIDBCAS	180
QY	181	GKVICPYNRRCVNTHGYSYCKCHTIGFELQYISGRYDCIDINECTMDSHTCSHHANCNTQ	240
Db	181	GKVICPYNRRCVNTHGYSYCKCHTIGFELQYISGRYDCIDINECTMDSHTCSHHANCNTQ	240
QY	241	GSPFKCKOGKKGNGLRCSAIPENSYSVEVLARAPETIDRIKILLAHNSMKKKAKIKNVY	300
Db	241	GSPFKCKOGKKGNGLRCSAIPENSYSVEVLARAPETIDRIKILLAHNSMKKKAKIKNVY	300
QY	301	PEPTRPTPKVNILOPFNYEETIVSSGSGSHG	331
Db	301	PEPTRPTPKVNILOPFNYEETIVSSGSGSHG	331

CC agents for regulating cellular processes like asthma, graft
 CC versus-host diseases, rheumatoid arthritis, psoriasis, inflammatory
 CC bowel disease, septic shock, ulcerative colitis, Crohn's disease,
 CC chronic myelogenous leukemia, cancer, liver disease, Hodgkin's
 CC disease, osteoarthritis, Lyme's disease, cachexia and autoimmune
 CC diseases e.g. myasthenia gravis, autoimmune diabetes and systemic
 CC lupus erythematosus. The nucleic acids are also useful for producing
 CC transgenic animals and the TANGO polypeptides themselves. Partial
 CC TANGO-128, 140, 197, 212, 213, 224, 239 sequences are useful in
 CC forensic biology, for diagnostic assays, prognostic assays,
 CC pharmacogenomics and for monitoring clinical trials. TANGO
 CC polypeptides are suitable for both prophylactic and therapeutic
 CC methods for treating a subject at risk of a disorder or having a
 CC disorder associated with aberrant TANGO expression. A wide range
 CC of cellular disorders can be treated.

XX Sequence 333 AA;

SO Query Match 48.7%; Score 1491; DB 21; Length 333;
 Best Local Similarity 76.2%; Pred. No. 5,2e-102;
 Matches 260; Conservative 27; Mismatches 46; Indels 8; Gaps 3;

QY 1 MRLPMSLALPRLLSWVAGFGNMAASARHHGLASAROPGVCHYGTKLACCYGMRNSKGY 60
 Db 1 MOPWGLALPRLLPVAVGAGV--TSPMDYGLSALAHOGVCHYGTAKACCYGMRNNKGV 58
 QY 61 CEATCEPCRCFGEVGPVKRCRCFPGYTGTCTCSQDVNECGMKPRPCQHRVCVNTHSGYKFC 120
 Db 59 CEAMCEPCRCFGEVGPVKRCRCFPGYTGTCTCTODVNECGVAFRCQHRVCVNTHSGYKFC 118
 QY 121 LSGHMLPMDATCVNSRTCAAMINCOYSCDTEBGPCLCPSSGLRLAPNGRCLDIDECAS 180
 Db 119 LSGHMLPMDATCVNSRTCAAMINCOYSCDTEBGPCLCPSSGLRLAPNGRCLDIDECAS 178
 QY 181 GKVICPYNRCVNTFGSYCKHIGFELQYISGRYDCIDINECTDMSHTSCSHHANCFTQ 240
 Db 179 SKAVCPNSNRRCVNTFGSYCKHIGFELQYISGRYDCIDINECTDMSHTSCSHHANCFTQ 238
 QY 241 GSFCKCKOGYKGNGLRCSAIPENSVEKLEAPGTIKDIRKLLAHKMSKKAKIKKNT 300
 Db 239 GSFCKCKOGYKGNGLRCSAIPENSVEKLEAPGTIKDIRKLLAHKMTKKKVKALKNT 298
 QY 301 PEPTPTPKVNLQPFNTEELVSRGNSHGSKKNEEKAKE 341
 Db 299 PRPASTRPVKVNL--PYSESEGVSRGRNYDG-----EQKKKK 333

RESULT 14
 AAY76110
 ID AAY76110 standard; Protein: 284 AA.
 AC AAY76110;
 XX 27-MAR-2000 (first entry)
 DT Rat TGF-beta homologue, SEQ ID NO:389.
 DE
 XX Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
 KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
 KW secreted; transmembrane; inflammation; cancer; neurological disease;
 KW angiogenesis; tumour vascularisation; growth disorder;
 KW developmental disorder; skin wound; hair follicle disorder;
 KW anti-inflammatory; cytostatic; neuroprotective; vulnerey.

OS Rattus sp.
 XX
 XX W09955865-A1.
 PN 04-NOV-1999.
 PD 29-APR-1999; 99WO-NZ00051.
 PF 29-APR-1998; 98US-0069726.
 XX PR

PR 09-NOV-1998; 98US-0188930.
 XX (GENE-) GENESIS RES & DEV CORP LTD.
 PA Strachan L, Sleeman M, Watson JD, Onrust R, Kumbje A, Murison JG;
 XX WPI: 2000-072177/06.
 DR N-PSDB; AAZ61818.
 XX
 PT Novel polynucleotides useful for the treatment of various conditions
 PT including wounds and cancer -
 PS Claim 4; Page 220-221; 235pp; English.
 XX The invention relates to novel nucleic acid sequences derived from rat
 CC dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
 CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying
 CC cells. Polypeptides of the invention may be used to treat inflammation,
 CC cancer and neurological diseases. The proteins may be used to stimulate
 CC the growth and motility of keratinocytes, to inhibit the growth of
 CC cancer cells, to modulate angiogenesis and tumour vascularisation, to
 CC modulate skin inflammation, to modulate epithelial cell growth and to
 CC inhibit binding of HIV-1 to leukocytes. The invention may also be used
 CC to treat growth and developmental defects, skin wounds and hair follicle
 CC disorders. Sequences AAY75942-Y76123 represent polypeptides encoded
 CC by cDNA sequences derived from several mouse, rat or human skin cell
 CC types. Sequences AAY75942-Y75947, AAY76020-Y76021, AAY76094-Y76104 and
 CC AAY76119 are proteins with an N-terminal signal sequence, indicating
 CC that they are secreted. Sequences AAY75986-Y75989, AAY76061-Y76071,
 CC AAY76106-Y76109 and AAY76121-Y76122 are proteins with one or more
 CC putative transmembrane domains.

XX Sequence 284 AA;

SO Query Match 42.1%; Score 1288; DB 21; Length 284;
 Best Local Similarity 82.6%; Pred. No. 4e-87;
 Matches 218; Conservative 14; Mismatches 30; Indels 2; Gaps 1;

QY 1 MRLPMSLALPRLLSWVAGFGNMAASARHHGLASAROPGVCHYGTKLACCYGMRNSKGY 60
 Db 23 MOPWGLALPRLLPVAVGAGV--TSPMDYGLSALAHOGVCHYGTAKACCYGMRNNKGV 80
 QY 61 CEATCEPCRCFGEVGPVKRCRCFPGYTGTCTCSQDVNECGMKPRPCQHRVCVNTHSGYKFC 120
 Db 81 CEAMCEPCRCFGEVGPVKRCRCFPGYTGTCTCTODVNECGVAFRCQHRVCVNTHSGYKFC 140
 QY 121 LSGHMLPMDATCVNSRTCAAMINCOYSCDTEBGPCLCPSSGLRLAPNGRCLDIDECAS 180
 Db 141 LSGHMLPMDATCVNSRTCAAMINCOYSCDTEBGPCLCPSSGLRLAPNGRCLDIDECAS 200
 QY 181 GKVICPYNRCVNTFGSYCKHIGFELQYISGRYDCIDINECTDMSHTSCSHHANCFTQ 240
 Db 201 SKAVCPNSNRRCVNTFGSYCKHIGFELQYISGRYDCIDINECTDMSHTSCSHHANCFTQ 260
 QY 241 GSFCKCKOGYKGNGLRCSAIPEN 264
 Db 261 GSFCKCKOGYKGNGLRCSAIPENH 284

RESULT 15
 AAB56049
 ID AAB56049 standard; Protein: 284 AA.
 AC AAB56049;
 XX 08-MAR-2001 (first entry)
 DT Skin cell protein, SEQ ID NO: 389.
 DE
 XX Rat; skin cell; cytostatic; anti-inflammatory; anti-HIV;
 KW neurologic; neuroprotective; vulnerey; immunomodulatory; vaccine;
 KW keratinocyte growth stimulation; cancer; angiogenesis inhibition;
 KW inflammation; neurological disease.

```

XX      Rattus sp.
OS
XX      WO200069884-A2.
XX      23-NOV-2000.
XX      15-MAY-2000; 2000WO-NZ00075.
XX      14-MAY-1999; 99US-0312283.
XX      (GENE-) GENESIS RES & DEV CORP LTD.
XX      Watson JD, Strachan L, Onrust R, Sleeman M, Kumble KD, Murison JG;
XX      WPI: 2001-007495/01.
XX      DR      N-PSDB; AAC99751.
XX
XX      New isolated polynucleotide used in the identification of genetic
XX      disorders and encoding polypeptides used for treating inflammatory
XX      disease, cancer and neurological diseases -
XX
XX      Claim 4; Page 295-296; 352pp; English.
XX
XX      The present sequence is a polypeptide which is expressed in
XX      mammalian skin cells. The polypeptide is useful for stimulating
XX      keratinocyte growth and motility, inhibiting the growth of cancer cells,
XX      modulating angiogenesis, inhibiting angiogenesis and vascularisation of
XX      tumours, modulating skin inflammation, stimulating the growth of
XX      epithelial cells, inhibiting the binding of human immunodeficiency virus
XX      (HIV)-1 to leukocytes, and treating inflammatory disease, cancer and
XX      neurological diseases. The polynucleotide can be used as a marker, in
XX      the identification of genetic disorders, and for the design of
XX      oligonucleotides for examining expression patterns.
XX
XX      .SQ      Sequence      284 AA;
XX
XX      Query Match      42.1%; Score 1288; DB 22; Length 284;
XX      Best Local Similarity      82.6%; Pred. No. 4e-87;
XX      Matches 218; Conservative 14; Mismatches 30; Indels 2; Gaps 1;
XX
QY      1      MPLPSLALPILLISWAGGFGMAASARHGLASAROPGVCHYGTKLACCYGMRNRSGV 60
QY      23      MOPFGLALPLLPVAVAGVG--TSPRDYWLPAALAHQPEVCHYGTAKCYGMRNRSGV 80
DB
QY      61      CEATCEPGCKFEGECVGPNNKRCFPGYTGKTCSDVNECGMKPRPCQHRVCVNTGSKCFC 120
DB      81      CEAVCEPRCKFEGECVGPNNKRCFPGYTGKTCSDVNECGMKPRPCQHRVCVNTGSKCFC 140
QY      121     LSGHMLMDATCVNSRTCAITNCOYSCEPTEBGPCLCPSSGLRLAPNGRDCIDIDECCAS 180
DB      141     LSGHMLMDATCVNSRTCAITNCOYSCEPTEBGPCLCPSSGLRLAPNGRDCIDIDECCAS 200
QY      181     GKVICPYNRRCVNTFGSYKCHIFELQYISGRYDCIDINECTMDSHSCSHHANCFTNQ 240
DB      201     SKAVCPNRCVNTFGSYKCHIFELQYISGRYDCIDINECTMDSHSCSHHANCFTNQ 260
QY      241     GSFCKCKQGYKGNGLRCSAIPEN 264
DB      261     GSFCKCKQGYKGNGLRCSAIPEN 284
DB
```

Search completed: October 17, 2002, 16:00:23
Job time : 39 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2002, 15:59:49 ; Search time 15 Seconds
(without alignments)
900.490 Million cell updates/sec

Title: US-09-687-860-24
Perfect score: 3060
Sequence: 1 MPEPMSLALPLLSWVAGF.....VDCVLVSGICPPSLLSVD 553

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

Issued_patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCtus.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	354.5	11.6	1833	US-08-479-722B-2	Sequence 2, Appl1
2	354.5	11.6	1833	PCT-US95-02251-18	Sequence 18, Appl1
3	339.5	11.1	956	US-08-897-443-3	Sequence 3, Appl1
4	335.5	11.0	2703	US-08-185-432-19	Sequence 19, Appl1
5	332.5	10.5	886	US-09-110-116-3	Sequence 3, Appl1
6	332.5	10.5	2556	US-08-083-590A-20	Sequence 20, Appl1
7	322.5	10.5	2556	US-08-532-384-20	Sequence 20, Appl1
8	320.5	10.5	638	US-08-897-443-1	Sequence 1, Appl1
9	320.5	10.5	2556	US-08-185-432-17	Sequence 17, Appl1
10	320.5	10.5	1394	517197-30	Patent No. 517197
11	313	10.2	443	US-08-833-963C-2	Sequence 2, Appl1
12	313	10.2	443	US-08-980-514-1	Sequence 1, Appl1
13	313	10.2	448	US-08-884-072-1	Sequence 1, Appl1
14	313	10.2	448	US-09-212-168-1	Sequence 1, Appl1
15	311.5	10.2	2523	US-08-185-432-18	Sequence 18, Appl1
16	310.5	10.1	810	US-08-820-170A-34	Sequence 34, Appl1
17	310.5	10.1	810	US-09-055-699-34	Sequence 34, Appl1
18	310.5	10.1	810	US-09-273-565-34	Sequence 34, Appl1
19	310.5	10.1	810	US-09-565-538-34	Sequence 34, Appl1
20	297.5	9.7	1218	US-09-214-278-7	Sequence 7, Appl1
21	296.5	9.7	1010	US-08-882-046-7	Sequence 7, Appl1
22	296.5	9.7	1036	US-09-068-740A-6	Sequence 6, Appl1
23	296.5	9.7	1187	US-09-068-740A-7	Sequence 7, Appl1
24	296.5	9.7	1218	US-08-400-159-6	Sequence 6, Appl1
25	296.5	9.7	1218	US-08-611-729A-6	Sequence 6, Appl1
26	296.5	9.7	1218	US-08-882-046-2	Sequence 2, Appl1
27	296.5	9.7	1218	US-09-068-740A-11	Sequence 11, Appl1

28	296.5	9.7	1219	US-08-882-046-5	Sequence 5, Appl1
29	296	9.7	2471	US-08-185-432-16	Sequence 16, Appl1
30	296	9.7	2471	US-08-083-590A-19	Sequence 19, Appl1
31	296	9.7	2471	US-08-532-384-19	Sequence 19, Appl1
32	293	9.6	337	US-09-188-930-186	Sequence 186, App
33	292	9.5	713	US-08-872-855-5	Sequence 5, Appl1
34	291	9.5	387	US-08-884-072-5	Sequence 5, Appl1
35	291	9.5	387	US-08-833-963C-9	Sequence 9, Appl1
36	291	9.5	387	US-08-980-514-3	Sequence 3, Appl1
37	291	9.5	387	US-09-212-168-5	Sequence 5, Appl1
38	291	9.5	652	US-08-751-305-2	Sequence 2, Appl1
39	288.5	9.4	1193	US-08-400-159-10	Sequence 10, Appl1
40	288.5	9.4	1193	US-08-611-729A-10	Sequence 10, Appl1
41	287.5	9.4	722	US-08-872-855-4	Sequence 4, Appl1
42	287.5	9.4	722	US-08-981-392-12	Sequence 12, Appl1
43	284.5	9.3	1248	US-08-882-046-6	Sequence 6, Appl1
44	282	9.2	1065	US-08-400-159-8	Sequence 8, Appl1
45	282	9.2	1212	US-09-214-278-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1

US-08-479-722B-2
Sequence 2, Application US/08479722B

Patent No. 6074840

GENERAL INFORMATION:
APPLICANT: Bonadio, Jeffrey

APPLICANT: Yin, Mushan
TITLE OF INVENTION: LATENT TGF (BINDING PROTEIN (LTBP)

	Query Match	11.6%	Score 354.5:	DB 3;	Length 1833;	
	Best Local Similarity	33.6%;	Pred. No. 3e-20;			
	Matches	99;	Conservative	27;	Mismatches 92;	Indels 77; Gaps 19.
OY	37 QPGVCHYT-----KIACCYGRNRNSKGYCAETC---PG-CKFGECV--GPNK-	80				
Db	947 QPVGSGGRCSNTEGSSYNHCECDRGIMWAKKGCOBINECRHGHTCTPDGCACVSPGSYTCL	1006				
OY	81 RCFPYGTATKS-QDVNECGMKPRPCQH-RCVNTGSKYCFCFLSGHMLPD-----	129				
Db	1007 ACEEGYVGSGSCVDVNEEC-LTPGICTHGRRCINMGSSFSCSEPEGYEVPDKKKGDVDE	1065				
OY	130 -----ATCYVN---SRTCAMINQY-----SCEPTREGPOC----LQPS-----	160				
Db	1066 CASRASCPYGLCLNTEGSTSTCASCGSVWNEDTRACEDLDE---CAFPGVCPTGYCTINT	1122				
OY	161 -----SGRLAPNGRDCLDIDECASGVKVIPLYNNRNCVNTFGSYCKRHIGELQY	210				
Db	1123 VGSFSGCKDDGGYRPNPJGNRCEDVEDDECGPSQC-RGGECCNTEBSYQLCHOGFQL--	1179				
OY	211 ISGRVDCIDINICTMDSTHSKSHANCFMTOGSGFFKCKRKQGYG--NGLRCSAIP E 263					
Db	1180 VNGTM-CEDEVNCEVEEH-CAPHGECLNSLSGFLLCACAFSAABEGTTCQDVDE 1232					

```

1      RESULT 2
2      PCT-US95-02251-18
3      ; Sequence 18, Application PC/TUS9502251
4      ; GENERAL INFORMATION:
5      ; APPLICANT:
6      ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE
7      ; TITLE OF INVENTION: CELLS
8      ; NUMBER OF SEQUENCES: 18
9      ; CORRESPONDENCE ADDRESS:
10     ; ADDRESSEE: Arnold, White & Durkee
11     ; STREET: P. O. Box 4433
12     ; CITY: Houston
13     ; STATE: Texas
14     ; COUNTRY: United States of America
15     ; ZIP: 77210
16     ; COMPUTER READABLE FORM:
17     ; MEDIUM TYPE: Floppy disk
18     ; COMPUTER: IBM PC compatible
19     ; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
20     ; SOFTWARE: PatentIn Release #1.0, Version
21     ; SOFTWARE: #1.30
22     ; CURRENT APPLICATION DATA:
23     ; APPLICATION NUMBER: PCT/US95/02251
24     ; FILING DATE: CONCURRENTLY HEREWITH
25     ; CLASSIFICATION:
26     ; PRIOR APPLICATION DATA:
27     ; APPLICATION NUMBER: US 08/316,650
28     ; FILING DATE: 30-SEP-1994
29     ; CLASSIFICATION:
30     ; APPLICATION NUMBER: US 08/199,780
31     ; FILING DATE: 18-FEB-1994
32     ; CLASSIFICATION:
33     ; ATTORNEY/AGENT INFORMATION:
34     ; NAME: Parker, David L.
35     ; REGISTRATION NUMBER: 32,165
36     ; REFERENCE/DOCKET NUMBER: UMIC009P--
37     ; TELECOMMUNICATION INFORMATION:
38     ; TELEPHONE: (512) 418-3000
39     ; TELEFAX: (713) 789-2679
40     ; TELEX: 79-0924
41     ; INFORMATION FOR SEQ ID NO: 18:
42     ; SEQUENCE CHARACTERISTICS:
43     ; LENGTH: 1833 amino acids
44     ; TYPE: amino acid
45     ; TOPOLOGY: linear
46     ; MOLECULE TYPE: protein
47     ;
48     ; PCT-US95-02251-18

```

Query Match	11.6%: Score 354.5; DB 5; Length 1893;
Best Local Similarity	33.6%: Pred. No. 3e-20;
Matches 99; Conservative	27; Mismatches 92; Indels 77; Gaps 19;
QY 37 QPVCYHYGT-----KLCACCYGMRNRSKGVCEAYCE--PG-CKFGECV--GPNKC-80	
Db 947 QPVGVSGRGSCNTEGSSYHCECDREGIIMWRKGHODINICRHFGTCDPDGRCVIVSPGSYTCL-1006	
QY 81 RCFPEYTKTOS-QDVNRCGKMKPRFCQ-RCVNTGSKYKFCFLSGHMLMPD-----129	
Db 1007 ACEBEYVAGSSGSCVDVNNC-LTPGICHTGRCLNMGSTRSCSEPEPEYTPDKKGGRDVE-1065	
QY 130 -----ATCVN---SRTCAMINQY-----SCEDTEGGPQC-----LCPS-----160	
Db 1066 CASRASCTGLCLNTEGSGFTSACOSGYWVEDGTACEDLE--CAEPGVCPHYGTNT-1122	
QY 161 -----SGLRLAPAGRCGLDIDECASGKVICPYNRRCVNFGSYYYCKCHIGFEIYQ-210	
Db 1123 VGSFSCKDDOGGRPNPLGNRCEDVDECEBGPSSC-RGEGCKNTGSGYQCLCHQGFOL--1179	
QY 211 ISGRGDCIDINECTMSHTSCSHNACFPMQSGFKCKKQGYKG--NGRCSAIRE-263	
Db 1180 VNGTM-CEDEVNCEGGEH-CAHGECLMSLGSFFCLCAPGFSAGSGTRCDQDVE-1232	

```

1 RESULT 3
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
10
```

```

RESULT 5
US-09-110-116-3
: Sequence 3, Application US/09110116
: Patent No. 6013479
:
: GENERAL INFORMATION:
:
: APPLICANT: Xu, Hong
: APPLICANT: Cohen, Victoria L.
: APPLICANT: Stuart, Susan G.
:
: TITLE OF INVENTION: HUMAN EMRL-LIKE G PROTEIN COUPLED
:
: TITLE OF INVENTION: RECEPTOR
:
: FILE REFERENCE: PF-0550 US
:
: CURRENT APPLICATION NUMBER: US/09/110,116
:
: CURRENT FILING DATE: 1998-07-02
:
: NUMBER OF SEQ. ID NOS.: 4
:
: SOFTWARE: fastseq for Windows Version 3.0
:
: SEQ ID NO 3
:
: LENGTH: 886
:
: TYPE: PRT
:
: ORGANISM: HOMO SAPIENS

```

```

; FEATURE:
; OTHER INFORMATION: 784994, GenBank
US-09-110-116-3

```

Query Match	10.5%	Score 323.5;	DB 3,	Length 886;
Best Local Similarity	30.2%	Pred. No. 4.7e-18;		
Matches 98;	Conservative 29;	Mismatches 112;	Indels 85;	Gaps 19

```

QY 11 LLSFWAGGFCFNMAARHNGLLAARPGV-----CHYGTKLACCYGMRNRSGKVCATCE 66
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6 LLLFW-----GCCVMHSMGEHIRTFRKRNPTKGNKNCROST--LCPRAYATCTYTVDSYITCK 59
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 67 PG-----CK-FGEC-----VGN-----KCRCEPGYTGKT--- 90
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 60 QGFLSSNGQNHFKRPYRKCDIDECSSPQPCGPNSSCKMLSGRYKSCILDGFPSPGND 119
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 91 -----CSODVNECGMKPRPC--QHR--CVNTHGSGYKFCULSGHMLMDATCVNSRTC 138
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 120 WVPKPGNFST--DINEC--LTSRYCPEHSDCVNMSGYSQCVQG--FISNSTCEDVNEC 176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 139 AMINC-----QYSCDETEBERPQCLC---PSSGLRIAPNG--RDCLDIDECASGKVICPYNR 189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 177 ADPRACEPHATCNTVYVNSCFCPNPFEGSSGHILSCGLKASCEIDDECTE---MCPINS 233
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 190 RCVATPFSYTCCKAIGF-----ELQYISGRYCDLIDINECTMDSTCHHANCFTVQSFK 244
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 234 TCTMTPTPSYFCTCHPGFAPSSGGQINLFPDQGECHDIDECRODPSTCGPNSICTNALGSYS 293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 245 CKCKQGYKGNGLRCSAIPENSVKR 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 294 CGCTVGFHNP-----PEGSQKD 310
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 6
 US-08-083-590A-20
 ; Sequence 20. Application US/08083590A
 ; Patent No. 5786158
 ;
 GENERAL INFORMATION:
 APPLICANT: Artaavanis-Tsakonas, S. et al.
 TITLE OF INVENTION: Therapeutic And Diagnostic Methods
 TITLE OF INVENTION: And Compositions Based On No. 5786158cn Proteins And
 TITLE OF INVENTION: Nucleic Acids
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036
 ;
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/083,590A
 FILING DATE: 25-JUN-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Mistrock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 7326-015
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 790-9090
 TELEFAX: 212 8698864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2556 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown

MOLECULE TYPE: peptide
US-08-083-590A-20

Query Match	10.5%;	Score 322.5;	DB 1;	Length 2556;
Best Local Similarity	25.3%;	Pred. No. 1.9e-17;		
Matches 94;	Conservative 31;	Mismatches 108;	Indels 139;	Gaps 16;

[illegible]

RESULT 7
 US-08-532-384-20
 : Sequence 20, Application US/08532384
 : Patent No. 6083904
 : GENERAL INFORMATION:
 : APPLICANT: Arcaanis-Tsakonas, S. et al.
 : TITLE OF INVENTION: Therapeutic And Diagnostic Methods
 : TITLE OF INVENTION: And Compositions Based On No. 6083904ch Proteins And
 : TITLE OF INVENTION: Nucleic Acids
 : NUMBER OF SEQUENCES: 21
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Pennie & Edmonds
 : STREET: 1155 Avenue of the Americas
 : CITY: New York
 : STATE: New York
 : COUNTRY: U.S.A.
 : ZIP: 10036
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/532,384
 : FILING DATE:
 : CLASSIFICATION: 424
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/083,590
 : FILING DATE: 25-JUN-1993
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Mistock, S. Leslie
 : REGISTRATION NUMBER: 18,872
 : REFERENCE/DOCKET NUMBER: 7336-015
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 212 790-9090
 : TELEFAX: 212 8698864/9741
 : TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 2556 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-532-384-20

Query Match 10.5%; Score 322.5; DB 3; Length 2556;
Best Local Similarity 25.3%; Pred. No. 1.9e-17;
Matches 94; Conservative 31; Mismatches 108; Indels 139; Gaps 16;

36 ROPGVCHVGR-----KLACCYGMRNRKVCPE---ATCEPCCKFGECV---GPMKCRCP 84
112 RNSGTCDLTLTEYKCKRCPRMSGKS---COADPCASNPANGGCLPFLASTYICHP 168
85 GYTKCTSDQVNECGMKPRPCQH--RCVNTGSKYKFCCLSGH----- 124
169 SFHGPTCRQDVNECGOKPRLCRHGTCCHNEVGYRCVCRATHNTPNCERPYVPCSPPCQ 228
125 -----MLMPDPT-----CVNRTCA-----MNC----- 143
229 NGTCRPTGDTHEACLPGETGQNCENIDDCPGNNCKNGACVQDVNTYNCPCPEPWT 288
144 -QYSCDETEE-----GPQCLPSSSGLRLAPNGRDCLD-IDECASGK 182
289 GQYTEVDDEGQALPNNACONGTCHNTHGYNVCVNW-----TGECSSNIDDCAS-- 341
183 VICPYNRCVNTFGSYCKCHIG-----FELYISGRYDCI- 218
342 AACFHGATGCHDRVASFYCECPHGRTGLCHLDACISNPNCGNSCDTNPVNGRAICTCP 401
219 -----DINCTMDSHSCSHANCFNTOGSEKCKCKGQYKNGNIGRCAIENSUYKE 268
402 SGYTGPCASQDVDECSIGANPCBHAGKCNITLGSFECOCLOGY--TGRCEIDVNECVSN 459
269 VLRAPTIKDRI 280
460 PCQNDATCIDLQI 471

RESULT 8
US-08-897-443-1
Sequence 1, Application US/08897443
Patent No. 5981263
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Puri
APPLICANT: Kaser, Mathew
TITLE OF INVENTION: HUMAN MATRILIN-3
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/897,443
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0348 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 638 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: UTRSN0702
CLONE: 681719
US-08-897-443-1

Query Match 10.5%; Score 320.5; DB 2; Length 638;
Best Local Similarity 34.3%; Pred. No. 4.4e-18;
Matches 80; Conservative 27; Mismatches 91; Indels 35; Gaps 14;

38 PG--VCHGTGLACCYGRNRN-----KGYC---EATCEPCCKFGECVGPNNKCRCPG 85
257 PGSYVCR-----CKQGYILNSDQTTCTRIODLCAMEDHNCQLCV--NVPGSFVCCQCYSG 308
86 YT-----GKTCSDQVNECGMKPRPCQHRVNTGSKYKFCCLSGHMLMPD-ATCVNSRTCA 140
309 YALMEDGKRCVA-VDYCASEMHGCEHCYNADGSLTLCHEGFALNPDKTCTYIDYAS 367
141 IN--QYSCDETEBGPQCLPSSSGLRLAPNGRDCLDIDECASGVYICPYNRCVNTFGSY 198
368 SNHGCGHCEVNTDYSYCHC-LKGFTLNPDKTKCRIRINYCALNRPGEH--ECVNMESY 424
199 YCKCHIGELYIGRYCIDINCTMDSHSCSHANCFNTOGSEKCKCKGQY 251
425 YCRCHRGYTL-DPNKG--TCSRVHDCAQDHCCEQ--LCLNTEDSFVCCSGE 473

RESULT 9
US-08-185-432-17
Sequence 17, Application US/08185432
Patent No. 5750652
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, Spyridon
APPLICANT: Bussseau, Isabelle
APPLICANT: Diegerich, Robert J.
APPLICANT: Xu, Tian
APPLICANT: Matsuno, Kenji
TITLE OF INVENTION: DELTAEX PROTEINS, NUCLEIC ACIDS, AND
TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,432
FILING DATE: 21-JAN-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2556 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-185-432-17

Query Match 10.5%; Score 320.5; DB 1; Length 2556;
 Best Local Similarity 25.38; Pred. No. 2.8e-17;
 Matches 94; Conservative 31; Mismatches 108; Indels 139; Gaps 16;

36 ROPGVCHYGT-----KIACCYGMRNRKSGVCE-----ATCEPCCKFGEYV---GPNKCRCFP 84
 112 RNGTCDLTLTLEKCKRCPKPMGSKS---CGADPCASNPANGGCGCLPFKASYICHCPP 168
 85 GYTKCTSQDVNECGMKRPPCOH--RCVNTGSKYKCFCLSGH----- 124
 169 SFHGPCTQWQDVNECGOKPRLCRHGGTCHNEVGYRCVCRATHTGPNCEMPYVPCSPQCQ 228
 125 -----MLMPDAT-----CVNRTCA-----MINC----- 143
 229 NGCTCRPTGDTHECACLPGFTGQNCENIDCGEKNCKNGACVGYNTYNCPCPPPMT 288
 144 -QYSCDETEE-----GPQCLCPSSGLRLAPNGRDCLD-IDECASGK 182
 289 GGYCTEDVDEQGLMPNACQNGGTCHNTHGYNVCVNGM-----TGEDCSEIIDCAS-- 341
 183 VICVYRNCVNTFGSYKCHIG-----FELYISGRIDCI-- 218
 342 AACFHGATCHDRVASFCECPHGRTGLCHLNDACISNPNEGSNCDTNPVNGKAITCPC 401
 219 -----DINECTMDSHNCFNTQGSFCKCKQGYKNGGLRCSAIPENSYKE 268
 402 SGYTPRACSDVDECSLIGANPCENHAGKINILGSECCOLOGY--TGPRCELDIVNECVSN 459
 269 VLRAPGITKRI 280
 460 PCQNDATCLDQI 471

RESULT 10
 517197-30
 Patent No. 5177197
 APPLICANT: KANZAKI, TETSUO; OLOFSSON, ANDERS; MOREN, ANITA;
 WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSON-WELSH,
 LENA; HELDIN, CARL-HENRIK
 TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING
 HUMAN TRANSFORMING GROWTH FACTOR-BETA1-BINDING PROTEIN
 NUMBER OF SEQUENCES: 53
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/487,343
 FILING DATE: 27-FEB-1990
 SEQ ID NO: 30:
 LENGTH: 1394
 517197-30

Query Match 10.5%; Score 320; DB 6; Length 1394;
 Best Local Similarity 31.28; Pred. No. 1.4e-17;
 Matches 89; Conservative 23; Mismatches 93; Indels 80; Gaps 16;

41 CHYGTKLACGWMNRNSKGC-----ATC-----EPG-----CKFG----- 72
 654 CEY-----CDSGRYMTQGRCEIDECINPSTCPDECVNPGSYQVCPTEGFGFGWNGQ 708
 73 -----ECVGR-----KRCFPGYT-----GKTCSDVYVNCGMKRPCCQHRC 109
 709 CLVDDECLPFPVNCANGDCSNLEGSYMSCGHGYRTDPHKHC-RDIDECQOGNLCVNGQC 767
 110 VNTGSGYKCFCLSGHMLMPDA-----TCVNSRTCAMINQYSCDETEEBPOCLCPSSG 162

Db 768 KNTGSEFRCTCGGGOYLSAANDOCEDIDECORHLICA-----HGOCNTEGSPQVC--DOG 822
 163 LRIAPNGRCDIDIECASCAGYICPNRRCVMTFGSYCKHIGFELYISGRYDCIDINE 222
 823 YRASGLDGHCEIDINECLDKSYC-QRGDCINTAGSYDCTCTPDGROL---DNKTKQDINE 878
 223 CTMDSH--TCSHMANCFNTQGSFKCKCKQGY--KNGILRCSAIPK 263
 879 C---EHPGLCGPQGECLNTEGSEFHCVCQGGFSISADGRTCEDIDE 920

RESULT 11
 US-08-833-963C-2
 Sequence 2, Application US/08833963C
 Patent No. 5916769

GENERAL INFORMATION:
 APPLICANT: Olsen, et al.
 TITLE OF INVENTION: Extracellular/bpidermal Growth Factor
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Ave
 CITY: Rockville
 STATE: MD
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patentln Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/833,963C
 FILING DATE: 11-Apr-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/US96/05033
 FILING DATE: 10-Apr-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Brookes, A. Anders
 REGISTRATION NUMBER: 36,373
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 301-309-8439
 TELEFAX: 301-309-8439
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 443 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-833-963C-2

Query Match 10.2%; Score 313; DB 2; Length 443;
 Best Local Similarity 35.88; Pred. No. 1.1e-17;
 Matches 69; Conservative 28; Mismatches 78; Indels 18; Gaps 10;

77 PNKCRCPFGY--TGKTCSDVNECGMKRPPCO--HRCVNTGSKYKCFCLSGH-MLMPDAT 131
 106 PNPCC--PGYERDDDDSCVDVDECAQALHDCRPSDDCHNLGSGYQCTCPDQYRKIGPE-- 161
 132 CVNSRTCAMINQYSCDETEBGPCLCPSSGLRLAPNGRDCLDIECASCAGYICPNRRC 191
 162 CVDIDECRYRCQHCRCVNLPGSFRCQC-EPGFOLGPNRRSCVDVNECDMG--APCEQRC 217
 192 VNTGSGYKCHHIGFELYISGRYDCIDINECTMDSHNCFNTQGSFKCKCKQGY 251
 218 FNSGYTFLCRHQGYELH--RDGFSCLDIDECSSSYSLQY--RCVNBPGRFSCHCPQGY 273
 252 KNGILR-CSAIPK 263
 274 QLATRLCQDIDE 286

RESULT 12

US-08-980-514-1

; Sequence 1, Application US/08980514

; Patent No. 6004753

; GENERAL INFORMATION:

; APPLICANT: Yue, Henry

; APPLICANT: Guegler, Karl J.

; APPLICANT: Shah, Puri

; TITLE OF INVENTION: HUMAN SI-5-ECMP-LIKE PROT

; TITLE OF INVENTION: EIN

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Dr.

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/980,514

; FILING DATE: Filed Herewith

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0436 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-855-0555

; TELEFAX: 650-845-4166

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 443 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: BRSTN013

; CLONE: 2786449

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

; US-08-980-514-1

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2002, 15:57:58 : Search time 22 Seconds

(without alignments)
2415.336 Million cell updates/sec

Title: US-09-687-860-24

Perfect score: 3060

Sequence: 1 MFLPMSIALPLLISWVAGF.....VDGVLLVSGLCPPDSLLSYDD 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_71:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3016	98.6	558	2 T17324	hypothetical prote
2	410.5	13.4	2907	2 A57278	fibriillin-2 precu
3	409.5	13.4	2918	2 A54105	fibriillin-2 precu
4	387.5	12.7	2871	2 A55567	fibriillin-1 - bovi
5	385.5	12.6	2871	2 A55567	fibriillin-1 precu
6	383.5	12.5	3002	2 A47221	fibriillin-1 precu
7	382	12.5	1221	2 A49457	fibriillin-2 precu
8	365.5	11.9	1574	2 T13954	MGF6 protein - ra
9	362.5	11.8	1184	2 A55184	fibriillin-2 precu
10	353.5	11.6	1620	2 T27283	hypothetical prote
11	350	11.4	683	2 C36346	fibriillin 1 precu
12	348	11.4	685	2 S78040	fibriillin 1 splice
13	338	11.0	601	2 B36346	fibriillin 1 precu
14	335.5	11.0	2703	2 A24420	notch protein - fr
15	335	10.9	705	2 S34968	fibriillin, splice
16	332	10.8	2437	2 S42612	transmembrane prot
17	331	10.8	589	2 T43210	fibriillin-ID precu
18	331	10.8	689	2 T42760	fibriillin, splice
19	330	10.8	712	2 T42990	fibriillin 1, splice
20	330	10.8	798	2 T22793	hypothetical prote
21	329	10.8	1712	2 A38261	masking protein
22	327.5	10.7	2318	2 S45306	notch3 protein - h
23	327.5	10.7	2321	2 S78549	notch3 protein - h
24	326	10.7	741	2 T46488	hypothetical prote
25	325.5	10.6	3507	2 T34513	latent transformin
26	325.5	10.6	1820	2 A55494	notch protein homo
27	325	10.5	2531	2 S18188	probable hormone r
28	322.5	10.5	886	2 A57172	notch protein homo
29	322.5	10.5	2555	2 A40043	notch protein homo

30	322	10.5	2531	2 A46019	Notch-1 protein -
31	320	10.5	1394	2 A35626	transforming growt
32	315	10.3	1964	2 T09059	notch4 - mouse
33	311	10.2	810	2 T10756	Notch-homolog protei
34	307.5	10.0	2471	2 A49128	cell-fate determin
35	303.5	9.9	2524	2 A35844	notch protein - Af
36	296.5	9.7	1220	2 A56136	jagged protein pre
37	294.5	9.6	2139	2 A35672	crumbs protein - f
38	291.5	9.5	1081	2 T31329	receptor tyrosine
39	291	9.5	387	2 T38449	extracellular prot
40	290	9.5	493	2 JC5621	epidermal growth f
41	289.5	9.5	2352	2 T30201	Notch homolog prot
42	287.5	9.4	722	2 T48324	DETTA-like 1 - mou
43	280	9.2	1106	2 T18739	hypothetical prote
44	279	9.1	1251	2 A57293	latent transformin
45	278	9.1	3623	2 T09456	Intrinsic factor-B

ALIGNMENTS

RESULT 1

T17324
hypothetical protein DKFZp564P2063.1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 15-Oct-1999 #sequence #revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T17324

R:Ducresthoef, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, September 1999

A:Reference number: 218727

A:Accession: T17324

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-558 <DUE>

A:Cross-references: EMBL:AL117610

A:Experimental source: fetal brain; clone DKFZp564P2063

C:Genetics:

A>Note: DKFZp564P2063.1

Query Match	Best Local Similarity	98.6%	Score 3016	DB 2	Length 558
Matches 547	Conservative	0	Mismatches 6	Indels 0	Gaps 0
QY 1	MFLPMSIALPLLISWVAGFGNAAARHGLASAROPGVCHYGTKLACCGMRNSGV	60			
DB 6	MFLPMSIALPLLISWVAGFGNAAARHGLASAROPGVCHYGTKLACCGMRNSGV	65			
QY 61	CEATCEPCCKFGECEVGPNNKRCRCPGYTGKTCSDVNECGMKRPRCQHRVNTGSKFCG	120			
DB 66	CEATCEPCCKFGECEVGPNNKRCRCPGYTGKTCSDVNECGMKRPRCQHRVNTGSKFCG	125			
QY 121	LSGHHMADATCVNSRTAMINCOYSCEDTEBGPCLPSSGLRLAPNGRCCLDIDECAS	180			
DB 126	LSGHHMADATCVNSRTAMINCOYSCEDTEBGPCLPSSGLRLAPNGRCCLDIDECAS	185			
QY 181	GVKICPYNRCVNFPGSYCKCHIGFELYTSGRDCIDINECTMDSTSHHACFNTQ	240			
DB 186	GVKICPYNRCVNFPGSYCKCHIGFELYTSGRDCIDINECTMDSTSHHACFNTQ	245			
QY 241	GSFKCKQGYKGNGLRCSAIPENSVEVLRAPTIKRIKLLAHKNSMKKAKIKNTV	300			
DB 246	GSFKCKQGYKGNGLRCSAIPENSVEVLRAPTIKRIKLLAHKNSMKKAKIKNTV	305			
QY 301	PEPTPTPPKYNLOPFNTEIYSGNSHGKKGKNEEMKBLEDEKREKALKNDIER	360			
DB 306	PEPTPTPPKYNLOPFNTEIYSGNSHGKKGKNEEMKBLEDEKREKALKNDIER	365			
QY 361	SLRGVFPFKVNEGEFFLLVORKALTSLEHKDLNLSVPCSFHNGICDMKODEDDFD	420			
DB 366	SLRGVFPFKVNEGEFFLLVORKALTSLEHKDLNLSVPCSFHNGICDMKODEDDFD	425			
QY 421	WNPADRDNAIGFYNAVPLAAGHKDGRILKLLPDLQPSNFCLLFDYRLAGDKVGLRV	480			

Query March	12.7%	Score 387.5	DB 2	Length 2871
Best Local Similarity	30.6%	Pred. No. 6.5	17	
Matches 89	Conservative 34	Mismatches 91	Indels 77	Gaps 10

QY	49	CCYGRNRNSKGVCEATCEPG-----CKEGECV--GPNKRC	82
DY	1081	CGRGCVTPDPDECKCEGYESGFMAMKNCMDIDECORDPLLCGCVLNTGSEYREC	1140
QY	83	FPGY--TGKTCSODVNECGMKPRPCQ--RCVNTGSKYCECLSGHMLMPDAT--CVNSRT	137
DB	1141	PRGQLANISACDINCELSAHLCPHRCVNLGLGKQACNPEYHSTPRLEFVDDIDE	1200
QY	138	CAMIN--CQYSEDTBEGPQCLPSSGLRLAPNGDCDIDDECASGKYIC-----PY	187
DB	1201	CSIMNGGCECTCTNSESYSCEC--QPGALMPDQSCDIDDECENPNICOGGQCTNIG	1259
QY	188	NRCV-----NFGSYCKCHIGELIYISR	214
DB	1260	EYRCICLYGFMASEDMKTCVDVNECDLNPNICLSGTCEINTGKSFCHCDMGYSK--RGK	1317
QY	215	YDCIDINCECTMDSHTCSHAANCFTQGSFKCKCKRGYVGNGLRGCAIPENS	265
DB	1318	TGCTDINCELGAHNCDDHAAVCTNTAGSFKCSGPMIGDGIKCTDLECS	1368

```

RESULT 5
A55624
Fibrillin-1 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 11-Jan-2000
C:Accession: A55624
J:Rlyn, W.; Smiley, E.; Germiller, J.; Sangunetti, C.; Lawton, T.; Perelra, L.; Ramirez,
J. Biol. Chem. 270, 1798-1806, 1995
A:Title: Primary structure and developmental expression of Fbn-1, the mouse fibrillin ge
A:Reference number: A55624; MUID:95130561
A:Accession: A55624
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2871 <YIN>
A:Cross-references: GB:L29454; NID:g575509; PIDN:AAA56040.1; PID:g575510
C:Genetics:
A:Gene: Fbn-1
C:Superfamily: unassigned EGF-related proteins; EGF homology
F:1201-1236/Domain: EGF homology <EGF>

Query Match          12.6%  Score 385.5;  DB: 2;  Length 2871;
Best Local Similarity 32.2%;  Pred. No. 8.8e-17;
Matches 86;  Conservative 34;  Mismatches 82;  Indels 65;  Gaps 11;

QY  39  GVCHGKTRKLACCGYRRRSKGVCEATCEPG-----CKRGECEV  75
      |||
Db  1127  GICH-----NTEGTYRCECPRHGLDPSINISACIDINECELSANLCPIGRCVN  1173
      |||

QY  76  --GPNKRCPEFY--TQKTSQDVNCEGMRPRCQHRVNTHTSGYKCFCLSGHMLMPDA  130
      |||
Db  1174  LIGKQCAQCNPPGHPHTHRRLCEVDIDECSIMNGCCFECTNSDGSYECSCQPGFALMPD  1232
      |||

QY  131  TCVNSRRTCAAMNCGYSCDTE--EGPQ-----CLCPSSGLRLAPNGNDICDIDEC  178
      |||
Db  1233  -----QRSTCID--QCEENPNICDGGOCTNIPGEYRLC-YDGFMASEDMMKTCVDVNEC  1284
      |||

QY  179  ASGVICVCPYNRCAVTFPSYCKCHIGFELQYISGRVPCIDINECTMASHSCSHANAFN  238
      |||
Db  1285  DLNPNIC-LISGCECENTKGSFLICHCMGYSGR--KKGTCGTIDINECEIGAHNCGRAVCTN  1341
      |||

QY  239  TQGFCKCKCKGQYKNGSLRCSAIPENS  265
      |||
Db  1342  TAGSGKSCSPGMWIDGICKCTDIDECs  1368
      |||

```

fibrillin 1 precursor - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 02-Jun-1995 #sequence-revision 25-Apr-1997 #text-change 21-Jul-2000

C:Accession: A47221, I54355, S17064, I59574, S17062, S62111, A34198

R:Corson, G.M.; Chalberg, S.C.; Dietz, H.C.; Charbonneau, N.L.; Sakai, L.Y.

Genomics 17, 476-484, 1993

A:Title: Fibrillin binds calcium and is coded by cDNAs that reveal a multidomain structure

A:Reference number: A47221, MUID:94010947

A:Accession: A47221

A:Molecule type: mRNA

A:Residues: 1-337, 'T', 339-1029 <COR>

A:Cross-references: GB:X63556

R:Peletre, L.V.; D'Alessio, M.; Ramirez, F.; Lynch, J.; Sykes, B.; Panglman, T.; Bo

Hum. Mol. Genet. 2, 961-968, 1993

A:Title: Genomic organization of the sequence coding for fibrillin, the defective gene

A:Reference number: I54355, MUID:93372860

A:Accession: I54355

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 132-3002 <PER>

A:Cross-references: GB:U13923, NID:g306745, PID:AB02036.1, PID:g306746

R:Maalen, C.L.; Corson, G.M.; Maddox, B.K.; Glanville, R.W.; Sakai, L.Y.

Nature 352, 334-337, 1991

A:Title: Partial sequence of a candidate gene for the Marfan syndrome.

A:Reference number: S17064, MUID:91304568

A:Accession: S17064

A:Molecule type: mRNA

A:Residues: 1030-3002 <MAS>

A:Cross-references: EMBL:X63556

R:Dietz, H.C.; Valle, D.; Francomano, C.A.; Kendzior, R.J.

Science 259, 680-683, 1993

A:Title: The skipping of constitutive exons in vivo induced by nonsense mutations.

A:Reference number: I59574, MUID:93157831

A:Accession: I59574

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 2217-2288, 'T', 2290-2325 <RES>

A:Cross-references: GB:S54426, NID:g264860, PID:AB25244.1, PID:g264861

R:Lee, B.; Godfrey, M.; Vitale, E.; Horl, H.; Mettel, M.G.; Saffarazi, M.; Tsipouras,

Nature 352, 330-334, 1991

A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two diff

A:Reference number: S17062, MUID:91304567

A:Accession: S17062

A:Molecule type: mRNA

A:Residues: 'VLTVVFIFLSTYKML', 944-1444 <LEE>

A:Cross-references: EMBL:X62008, NID:g31398, PID:CA56534.1, PID:g5924015

A:Accession: S62111

A:Molecule type: protein

A:Residues: 1166-1176, 'X', 1178-1180, 'D', 1182-1185 <LEE2>

R:Maddox, B.K.; Sakai, L.Y.; Keene, D.R.; Glanville, R.W.

J. Biol. Chem. 264, 21381-21385, 1989

A:Title: Connective tissue microfibrils. Isolation and characterization of three larg

A:Reference number: A34198, MUID:90070246

A:Accession: A34198

A:Molecule type: protein

A:Residues: 565-575, 1890-1892, 'T', 1894-1900 <MAD>

C:Comment: Fibrillin is a major component of elastin-associated microfibrils.

C:Genetics:

A:Gene: GDB:FBNI

A:Cross-references: GDB:127115, OMIM:134797, OMIM:154700

A:Map position: 15q21.1-15q21.1

A:Introns: 2236/1, 2258/1, 2297/1

C:Superfamily: unassigned EGF-related proteins; EGF homology

C:Keywords: alternative splicing; calcium binding; extracellular matrix; glycoprotein

F.1-3002/Product: fibrillin (5'-region exon A splice form) (fragment) #status predict

F.132-3002/Product: fibrillin (5'-region exon C splice form) #status predicted <MARC>

F.1332-1367/Domain: EGF homology <GEF>

F.1457-1492/Domain: EGF homology <GEF2>

F.2262-2295/Domain: EGF homology <GEF1>

RESULT 6	Query Match	12.54	Score 383.5	DB 2	Length 3002
447221	Best Local Similarity	32.68	Pred. No. 1.2e-16		
	Matches 87	Conservative 33	Mismatches 82	Indels 65	Gaps 11

QY 39 GVCHYGRKLACCGYWRNRSGVCEATCEPG-----CKFGEQCV- 75
 1258 GVCH-----NTEGSEYRCECPGHHQLSPNISACIDINECELSAHLCPGRGCVN 1304
 QY 76 --GNKCKCFPGY---TGKTSQDVNNEGCMKPRCQHHCVNTHGSKYKFCISGHHMLMDA 130
 1305 LIGRYOCACNPGYHSTPRLRCVLDIDECISINMGCEPCTSESGYECSCOPGFALMPD- 1363
 QY 131 TCVNSRTCAINCOYSCEDTE---EGPQ-----CLCPSSGLRLAPNGRDCLDIDEC 178
 1364 ----QRSTIDID---ECEDNPNICDGGCCTNIPGEYRLC-YDGFMASEDMKTCVDVNEC 1415
 QY 179 ASGRVICYPNRRCVNTEGSSYYCKHIGELQYISGRYDCIDINECTMDSHTCSHHANCFN 238
 1416 DLNPNIC-LSGTCENTKSFICHDGMYSGK--KGKGTCTDINECEICAHNCGKHAVCVN 1472
 QY 239 TQSEFKCKCKGKYGKNGLRCSAIPENS 265
 1473 TAGSFKCSGSPWIGDGIKCTDLDECS 1499

RESULT 7

A49457
 fibulin-2 precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 03-May-1994 #sequence.revision 03-May-1994 #text.change 11-Jan-2000
 C:Accession: A49457; S74095
 R:Pan, T.C.; Sasaki, T.; Zhang, R.Z.; Faessler, R.; Timpl, R.; Chu, M.L.
 J. Cell Biol. 123, 1269-1277, 1993
 A:Title: Structure and expression of fibulin-2, a novel extracellular matrix with
 A:Reference number: A49457; MUID:94064787
 A:Accession: A49457
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1221 <PAN>
 A:Cross-references: GB:X75285; NID:9437046; PID:CA53040.1; PID:9437047
 R:Sasaki, T.; Mann, K.; Murphy, G.; Chu, M.L.; Timpl, R.
 Eur. J. Biochem. 240, 427-434, 1996
 A:Title: Different susceptibilities of fibulin-1 and fibulin-2 to cleavage by matrix met
 A:Reference number: S74094; MUID:96439073
 A:Accession: S74095
 A:Molecule type: protein
 A:Residues: 226-238, 'X', 240-247, 260-275, 336-344, 'L', 346-361, 405-426, 566-568, 'EM', 569-586
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 C:Keywords: calcium binding; duplication; extracellular matrix; glycoprotein; homotrimer
 F:942-978/Domain: EGF homology <EG>

Query Match 12.5%; Score 382; DB 2; Length 1221;
 Best Local Similarity 28.5%; Pred. No. 6.3e-17;
 Matches 114; Conservative 39; Mismatches 139; Indels 108; Gaps 20;

QY 47 LACCYGWRNRSG---VCEATCEPG---CKFGE-CV---GPNKCRCPGYT---GKTCS 92
 878 LVCGRGYHANESEGCYVNECEGVHRCGSGQLCYNLPGSYRCKCKGFPGRDAGRRCI 937
 QY 93 QDVNECCGKPR-PCQHRCVNTHGSKYKFCISGHHMLMPAT--CVNSRTCAINCOYSCEDT 150
 938 -DVNECWVSPRLCQHTENPGRSCAAGFLAAGKCEDEVDNEETRRCSQECANI 996
 QY 151 EEGQCLCPSSGLRLAPNGRDCLDIDECASGR-VICPNRRCVNTEGSSYYCKHIGELQ 209
 997 YGSIQCYC-RQGYQLAEHGHTCTDIDECAGAGILCTE--RCVNVPGSYQACAPGQGYTM 1053
 QY 210 YISGRYDCIDINECTMDSHTCSHHANCFTQSGFKC---KCKQGY-KGNGLRCSAIPENS 265
 1054 MANOR-SEKDLDECALGTHNCSAETCHNIGSFRCLAFDPCPNVVRYSQTKCEHTTQD 1112
 QY 266 VKEVLRADGTIKDKIKLLAHKNSKKRAKIKNVTPETPTPTPKVNLDPNVEIEIVSRG 325
 1113 ITECOTSPARI---THYQLNFQTLGILVPAHIFRIGPAP----- 1147
 QY 326 GNSHGKKGKNGEKKKEGLEDEKREKALKNDIEKSLAGDYF---FPVNVNAGEF----- 377

Db 1148 -----AFAGDTISLTITGNEEGYFVTRRL 1172
 QY 378 ----GLIVORKALTSKLEHKDLNISVDCSFNHHGICDCKQ 413
 Db 1173 NAYTVGSVLQR-----SVLEPRDFALVDYMKL-----WRQ 1202

RESULT 8

T13954
 MEGF6 protein - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 20-Sep-1999 #sequence.revision 20-Sep-1999 #text.change 21-Jul-2000
 C:Accession: T13954
 R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
 Genomics 51, 27-34, 1998
 A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motif
 A:Reference number: Z14126; MUID:98360089
 A:Accession: T13954
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1574 <NAK>
 A:Cross-references: EMBL:AB011532; NID:93449293; PIDN:BA432462.1; PID:93449294
 A:Experimental source: strain Sprague-Dawley; brain
 C:Genetics:
 A:Gene: MEGF6

Query Match 11.9%; Score 365.5; DB 2; Length 1574;
 Best Local Similarity 36.8%; Pred. No. 9.4e-16;
 Matches 93; Conservative 25; Mismatches 96; Indels 39; Gaps 13;

QY 34 SARQGVCHYGTAKLACCGYWRNR-----NSKGYCEATCEPGCKFGEQCVGN 78
 92 SYROYATVARTVFRCCGWSQKPGQEGCLSDVDECAASNGCEGPC--C--NTVGGF 145
 QY 79 KCRCPGY-----TGKTSQDVNNEGCMKPRCQHRCVNTHGSKYKFCISGHHMLMPDA-TCV 133
 Db 146 YCRCPGYQLQDGGKTC-QDVDECAHNGGQHRCVNTPGSTLCECKGFRLLHDGRKCL 204
 QY 134 NSRTCAIN--COYSC-EDTEGPOCLCPSSGLRLAPNGRDCLDIDECASGRKVCIPYNR 190
 205 AISCSTLNGGCGQHCQVQLVYQHRCQCRPQ-YQLQEDGRCVRSPCAEGNGGCMH--I 261
 QY 191 CVNTPGSTLCECKHIGELQYISGRYDCIDINECTMDSHTCSHHANCFTQSGFKCKQ 250
 Db 262 COELNGLAHCHCHPQYQ--AADRTECDVDECAIGLAQCAH--GCLNTPGSSFKVCVCHAG 317
 QY 251 YK--GNGLRCSAI 261
 Db 318 YELGADGRQCYRI 330

RESULT 9

A55184
 fibulin-2 precursor - human
 N:Alternate names: protein DKFZ586A1519.1
 C:Species: Homo sapiens (man)
 C:Date: 27-Jan-1995 #sequence.revision 27-Jan-1995 #text.change 21-Jul-2000
 C:Accession: A55184; T08744
 R:Zhang, R.Z.; Pan, T.C.; Zhang, Z.Y.; Mattei, M.G.; Timpl, R.; Chu, M.L.
 Genomics 22, 425-430, 1994
 A:Title: Fibulin-2 (FBLN2): human cDNA sequence, mRNA expression, and mapping of the
 A:Reference number: A55184; MUID:95104855
 A:Accession: A55184
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1184 <ZHA>
 A:Cross-references: GB:X82494; NID:9575232; PIDN:CA57876.1; PID:9575233
 R:Mambuit, R.; Heudner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, March 1999
 A:Reference number: Z16471
 A:Accession: T08744
 A:Molecule type: mRNA

A:Residues: 655-719, /ODELTMAHCHCSRQFCVNTLGSFYCVNHYTLCADGILLNAHRCVD',720-853,'T',855-1140
A:Cross-references: EMBL:AL050095
A:Experimental source: adult uterus; clone DKF2p586a1519
C:Genetics:
A:Gene: GDB:FBLN2
A:Cross-references: GDB:293037; OMIM:135821
A:Map position: 3p25-3p24
A:Note: DKF2p586a1519.1
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: alternative splicing; extracellular matrix
F.1-27/Domain: signal sequence #status predicted <SIG>
F.28-1184/Product: fibulin-2 protein #status predicted <MAT>
:905-941/Domain: EGF homology <EGF>

Query Match	11.8%;	Score 362.5;	DB 2;	Length 1184;
Best Local Similarity	30.7%;	Pred. No. 1.1e-15;		
Matches 107;	Conservative 35;	Mismatches 149;	Indels 57;	Gaps 17;

```

QY      21 GNASARHGGLLASARPVCYHGT-----KLACCYGRNRNSK---VCEATCEPG 68
      || : || | | : | | |
Db      806 GNCVDINECTSLSEPCRGFSCINTVGSYTCQRNPILCARGHASDCAKVDVNECEETG 86

```

```

09      --CAGGECV----GPNRCRCFPGYT-----GKTCSDVDNNEGMKP-RCQQRHCYNTHSGSY 11
          | :         | :| :         | : |         | :||| |         |||
DbB    VHRCEGGVCVCHNLPGSYRCDCKAGFGORDAFGRGCI-DVNECMASPGRLCQHTCENTLGSY 92

```

QY 11/ KCFCISGHHMLNPD-TCVNSRTCAMINQCSCDEPHEEGPQCICPSSGRLEAPNRDCLDI 17
:
Db 925 RCSASGFLLADGKRCEDVNECEAQRCSQECANITYGYCIC-RGGTQLAEIDSGHTCTDI 98

```

170 DECAAGK-VLPYINRKNVTFGSYCKCHGFEIDYISGRDCIDINDCTMDSHCSHHA 23
    ||| |::|:|:| ||| |::| |::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 984 DECAAGILCTF--RLNLVPGSIQCACPCPGYLTMTANGR--SKDVLDECALGTNCSFAE 10

```

Db 1041 TCHNIQSFRLFECPNYYVQSKTCERTCHDFLECONSARI---THYQINFQTL 10

```

      | | : | | | : | |
Db 1098 LVPAHIFRIGPAPAFGTIALNI-----IKGNEE 1127

```

RESULT 10
T27283
hypothetical protein Y64G10A.f - *Caenorhabditis elegans*

C:\Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:\Accession: T27283
R:\Alnscough, R.

A:Reference number: Z20336
A:Accession: T27283
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Residues: 1-1620 <MIL>
A:Cross-references: EMBL:AL110498; NID:e1542303; PIDN:CAB54471.1; CESP:Y6
A:Experimental source: clone Y64G10A
C:Contact:

Query Match	11 69	Score 353.5	DB 2	Length 1620
A:gene: CESP:Y64G10A.1				
A:introns: 77/1; 116/1; 198/1; 282/1; 365/1; 425/1; 466/1; 548/1; 559/1;				

40 VCHGYMGLACCGWRNRNSKGYCF-----ATCF-----PGCGEGGCV---GPNK 79
 best local similarity 33.9%; Pred. NO. 5./e-15;
 Matches 82; Conservative 38; Mismatches 81; Indels 41; Gaps

Db 83 LCHNGG--TCVPSEHNDNEQVCECPVGFITGAKQDYDANEAMNNGGCEH--ECVNTIGTYY 13

Db 140 CRCHWGFELSDGNTCS - DIDECAVSNGGCSDRCVNSPGGFRCDPSDLYLHMDGRTCGK 19

[illegible]

RESULT 11
C36346

N/Alternate names: fibulin C
N/contains: fibulin 1 splice form A; fibulin 1 splice form C
C/Species: Homo sapiens (man)

C; Accession: C36346; A36346; A32826
R; Graves, W.S.; Tran, H.; Burgess, W.H.; Dickerson, K.
J. Cell Biol. 111, 3155-3164, 1990

A:Reference number: A36346; MUID:91100426
A:Accession: C36346
A:Molecule type: mRNA
A:Positives: 1-683 <AVC>

A:cross-references: GR:X3/43; NID:g31418; PIDN:CA3///2.1; PID:g31419
A:accession: A36346
A:molecule type: mRNA
A:Residues: 1-566 <AR2>

A; Title: Fibulin, a novel protein that interacts with the fibronectin receptor. Cell 58, 623-629, 1989

A:Accession: A32826
A:Molecule type: protein
A:Residues: 30-35, 'SX', 38-40, 'SH', 43-44 <R3>

A:Gene: GDB:FBLN1; FBLN
A:Cross-References: GDB:278285; OMIM:135820
A:Map position: 22q13.3-22q13.3

C;keywords: alternative splicing; glycoprotein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-683/Product: fibulin 1 splice form C #status predicted <MAT>

F:485-523/Domain: EGF homology <EGF1>
F:58,535,539/Binding site: carbohydrate (Asn) (covalent) #status predicted

Best Local Similarity 31.9%; Pred. No. 4.1e-15;
Matches 100; Conservative 37; Mismatches 120; Indels 56; Gaps

Db 341 CGRGYHINEGTRCVDVDECAAPPAEPGKGHRVNSPSGFRCECKTGYVFDGISPMC -VD 399

Db 400 VNECORYPGRLCGHKCENTLGSYLCSSGVGFRSLVDGRSCEDIIECCSSSPCSQECANVYG 459

Db 460 SYGCI-RRGYQLSDVDGVTCEIDIDEALPTGHHISY--RCINIPGSIQCSPSSGYRL 516

QY 209 QYISGRYDCIDINECTMDSHTCSHANCENFQSGFKC---KCRGGYKGN-GLRCSAIPEN 264

Db 517 A-PNCR-NCQDIDECVTGIIHNCISINIEICFNIGCAFRLCAFECEPNYRRSATRCERLPC 574

A:Molecule type: DNA
 A:Residues: 2504-2576,'E',2578-2611 <WHA2>
 C:Genetics:
 A:Gene: notch; opa
 A:Cross-references: FlyBase:FBgn0004647
 A:Map position: 8.96-9.36
 A:Introns: 53/3; 84/3; 171/3; 240/3; 283/3; 2333/3; 2436/3; 2588/3
 C:Superfamily: notch protein; tandem repeat; transmembrane protein
 C:Keywords: differentiation; tandem repeat; transmembrane protein
 F:21743/Domain: transmembrane #status predicted <TM1>
 F:237-328/Domain: EGF homology <EGF1>
 F:530-561/Domain: EGF homology <EGF1>
 F:568-599/Domain: EGF homology <EGF1>
 F:988-1019/Domain: EGF homology <EGF2>
 F:1064-1095/Domain: EGF homology <EGF3>
 F:1187-1218/Domain: EGF homology <EGF3>
 F:1746-1762/Domain: transmembrane #status predicted <TM2>
 F:1950-1982/Domain: ankyrin repeat homology <AN1>
 F:1983-2015/Domain: ankyrin repeat homology <AN2>
 F:1988-2004/Domain: transmembrane #status predicted <TM3>
 F:2017-2049/Domain: ankyrin repeat homology <AN3>
 F:2050-2082/Domain: ankyrin repeat homology <AN4>
 F:2083-2115/Domain: ankyrin repeat homology <AN5>
 F:2538-2568/Region: glutamine-rich
 F:2538-2568/Domain: neurogenic repetitive element #status predicted <OPA>

Query Match 11.0%; Score 335.5; DB 1; Length 2703;
 Best Local Similarity 26.4%; Pred. No. 1.4e-13;
 Matches 96; Conservative 34; Mismatches 106; Indels 127; Gaps 17;

QY 39 GVCHYGT-----KLACCYGM---RRNSKGVCEAT-CEPCCKFGEVCVPMK--CRCPGYTG 88
 Db 152 GTCLKLTLEETACANCANGYTERGETNKLCASSPCRNATCTALAGSSFTCSCPGFTG 211
 QY 89 KTGSDVNECGMKPRPCOH--RCVNTGSKYKFCISG-----HMLMDATCVNSRT 137
 Db 212 DTCSYDIEEC--QSNPKYGGTCVNTGHSYOCMPTGYTKDCDTKYPKSPSCQNGGI 269
 QY 138 CAMINCOYSCEDTE--EGPOC-----LCPSSG-----LRLAPN--GRDCL 173
 Db 270 CRNGLSTYCKCPKGFESKNCENYDCLGHLCONGGLCIDIGISDYTCRCPNNTGFRQ 329
 QY 174 -DIDECAS-----GKVICPYNRRCN 193
 Db 330 DDVDECAQRDHPVCQNGATCTNTHGYSCLICVNGMAGIDCSNNTDDCAACFGATCID 389
 QY 194 TFGSYTCKCHG-----FELYISGRDCT-----D 219
 Db 390 GVGSPFYCQCTKGTGLCHLDDACTSNPCHADAICTSPINGSYACSCATGYKGVDCSED 449
 QY 220 INECTMDSHTCSHANCENFTGSKFKCKQGYG-----NGLRCSAIPENSKEVLRAPG 274
 Db 450 IDECDQGS-PCENHICVNTPTGSTRCNCSGCTGPRCETNINECESHPCONEGSCLDDPG 508
 QY 275 TIK 277
 Db 509 TFR 511

RESULT 15
 S34968
 fibulin, splice form D precursor - mouse
 N:Alternate names: basement-membrane protein BM-90; calcium-binding protein BM-90
 C:Species: Mus musculus (house mouse)
 C>Date: 10-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 31-Jan-2000
 C:Accession: S34968; S36441; S13814
 R:Pan, T.C.; Kluge, M.; Zhang, R.Z.; Mayer, U.; Timpl, R.; Chu, M.T.
 Eur. J. Biochem. 215, 733-740, 1993
 A:Title: Sequence of extracellular mouse protein BM-90/fibulin and its calcium-dependent
 A:Reference number: S34968; MUID:93358897
 A:Accession: S34968
 A:Molecule type: mRNA
 A:Residues: 1-705 <PAN>

R:Pan, T.C.; Kluge, M.; Zhang, R.Z.; Mayer, U.; Timpl, R.; Chu, M.T.
 submitted to the EMBL Data Library, January 1993
 A:Description: Sequence of extracellular mouse protein BM-90/fibulin and its calcium-
 A:Reference number: S36440
 A:Accession: S36441
 A:Molecule type: mRNA
 A:Residues: 1-39,'P',41-705 <PAN>
 A:Cross-references: EMBL:X10854; NID:g396820; PIDN:CA50207.1; PID:g396821
 A:Experimental source: Cell-line F9 teratocarcinoma
 R:Kluge, M.; Mann, K.; Dziadek, M.; Timpl, R.
 Eur. J. Biochem. 193, 651-659, 1990
 A:Title: Characterization of a novel calcium-binding 90-kDa glycoprotein (BM-90) shar
 A:Reference number: S13814; MUID:91065369
 A:Accession: S13814

A:Molecule type: protein
 A:Residues: 28;31-49,'X',51-53,'XX',110-117,231-240,'X',242-243;339-362,'S',364-387,4
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 C:Keywords: alternative splicing; basement membrane; calcium binding; extracellular m
 F:1-29/Domain: signal sequence #status predicted <SIG>
 F:30-705/Product: fibulin, splice form D #status predicted <MAT>
 F:98,537,541/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.9%; Score 335; DB 2; Length 705;
 Best Local Similarity 34.1%; Pred. No. 3.9e-14;
 Matches 86; Conservative 31; Mismatches 87; Indels 48; Gaps 15;

QY 46 KLACCYGHRNSKGYC-----EATCEPCCKFGEVCVPMKCR-----CFPGY-----T 87
 Db 293 KLOCKSGFIIDALGNCIDINECLISAPCPYGCCTCINTGSGYTCOKNVPNCGRGYHNEE 352
 QY 88 KTGSDVNECGMKPRPC--OHRVNTGSKYKFCISGHLMDATCVNSRTCAMIN--- 142
 Db 353 GTRC-VDVDECAPRAEPKCGKHGCLNSPGRCECKAFTY--DGI---SRTCDINECQ 406
 QY 143 -----CQYSCDETEEGPOCLPSSGLRLAPNGRCLDIDECASGKYICPYNRRCVNTG 196
 Db 407 RYPRRLCGHKCNENTPGSPHCSC-SAGFRLSYDGRSCEDVNCILNS---PCSQECANYG 461
 QY 197 SYCKCHIGFELYISGRYDCLDINECTMDS--HTCSHANCFTGSKFKC-KQGYK- 252
 Db 462 SYQCYCRGYSGLSDVDG-VTCEIDIDECALPTGHICSY--RCINIPGSFQSCSPSSGYRL 518
 QY 253 -GNGLRCSAIP 263
 Db 519 APNGRNCDDIDE 530

Search completed: October 17, 2002, 16:01:56
 Job time : 32 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2002, 15:53:08 : Search time 14 Seconds
(without alignments)
1529.422 Million cell updates/sec

Title: US-09-687-860-24
Perfect score: 3060
Sequence: 1 MFLPMSLALPLLSSWAGF.....VDGVLLVSGLCPPSLSLSYDD 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	410.5	13.4	2907	1 FBN2_MOUSE	Q61555 mus musculus
2	409.5	13.4	2911	1 FBN2_HUMAN	P35556 homo sapien
3	387.5	12.7	2871	1 FBN1_BOVIN	P98133 bos taurus
4	385.5	12.6	2871	1 FBN1_MOUSE	Q61554 mus musculus
5	385.5	12.6	2871	1 FBN1_PIG	Q9TV36 sus scrofa
6	383.5	12.5	2871	1 FBN1_HUMAN	P35555 homo sapien
7	382	12.0	1221	1 FBN2_MOUSE	P37889 mus musculus
8	367	12.0	684	1 FBL1_CHICK	O73775 gallus gall
9	362.5	11.8	1184	1 FBL2_HUMAN	P98095 homo sapien
10	339	11.1	956	1 MTN2_MOUSE	Q08746 mus musculus
11	338	11.0	703	1 FBL1_HUMAN	P23142 homo sapien
12	335.5	11.0	2703	1 NOTC_DROME	P07207 drosophila
13	335	10.9	705	1 FBL1_MOUSE	Q08879 mus musculus
14	332	10.8	2437	1 NOTC_BRARE	P46530 brachydanio
15	331	10.8	956	1 MTN2_HUMAN	Q00339 homo sapien
16	330	10.8	712	1 FBL1_CAEEL	O77469 caenorhabd
17	329	10.8	1712	1 TGFB_RAT	Q00918 rattus norv
18	327.5	10.7	2318	1 MTN3_MOUSE	Q61982 mus musculus
19	325	10.6	2531	1 MTN1_RAT	Q07008 rattus norv
20	322.5	10.5	886	1 EMRI_HUMAN	Q14246 homo sapien
21	322.5	10.5	2444	1 MTN1_HUMAN	P46531 homo sapien
22	322	10.5	2531	1 MTN1_MOUSE	Q01705 mus musculus
23	320	10.5	1394	1 TGFB_HUMAN	P22064 homo sapien
24	315.5	10.3	1964	1 FBL1_MOUSE	Q9WV39 mus musculus
25	315	10.3	1964	1 MTN4_MOUSE	P31655 mus musculus
26	313	10.2	443	1 FBL4_HUMAN	Q95967 homo sapien
27	313	10.2	448	1 FBL5_HUMAN	Q9UB45 homo sapien
28	311.5	10.2	2524	1 NOTC_XENLA	P21783 xenopus lae
29	311	10.2	448	1 FBL5_RAT	Q9WV88 rattus norv
30	311	10.2	810	1 NEI1_RAT	Q62919 rattus norv
31	311	10.2	835	1 CD97_HUMAN	P48960 homo sapien
32	310.5	10.1	810	1 NEI1_HUMAN	Q92832 homo sapien
33	309.5	10.1	448	1 FBL5_MOUSE	Q9WVH9 mus musculus

34	308	10.1	443	1 FBL4_CRIGR	O55058 cricetus
35	294.5	9.6	2139	1 CRB_DROME	P10040 drosophila
36	292	9.5	714	1 DLI1_RAT	P97677 rattus norv
37	291	9.5	493	1 FBL3_HUMAN	Q12805 homo sapien
38	290.5	9.5	931	1 EMRI_MOUSE	Q61568 mus musculus
39	290	9.5	493	1 FBL3_RAT	Q35568 rattus norv
40	287.5	9.4	722	1 DLI1_MOUSE	Q61483 mus musculus
41	279.5	9.1	816	1 NEI2_MOUSE	Q62918 rattus norv
42	279	9.1	723	1 DLI1_RAT	Q00548 homo sapien
43	279	9.1	816	1 NEI2_HUMAN	Q90827 gallus gall
44	278	9.1	816	1 NEI2_MOUSE	Q61220 mus musculus
45	273	8.9	816	1 NEI2_HUMAN	Q99435 homo sapien

ALIGNMENTS

```

RESULT 1
FBN2_MOUSE STANDARD: PRT: 2907 AA.
AC 061555; 063957;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE Fibrillin 2 precursor.
GN FBN2 OR FBN-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9526370; PubMed=7744963;
RA Zhang H., Hu W., Ramirez F.;
RT "Developmental expression of fibrillin genes suggests heterogeneity
RT of extracellular microfibrils."
RL J. Cell Biol. 129:1165-1176(1995).
RN [2]
RP SEQUENCE OF 210-317 FROM N.A.
RX MEDLINE=94140368; PubMed=8307578;
RA Li X., Pereira L., Zhang H., Sanguineti C., Ramirez F., Bonadio J.,
RA Francke U.;
RT "Fibrillin genes map to regions of conserved mouse/human synteny on
RT mouse chromosomes 2 and 18."
RL Genomics 18:667-672(1993)
CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE
CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
CC -!- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING
CC EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L39790; AAA74908.1; -
CC EMBL: S69359; AAC60685.1; -
CC HSSP: P35555; IEMN.
CC
CC MGD: MGI:95490; Fbn2.
CC InterPro: IPR000152; Asx_hydroxyl.
CC InterPro: IPR000561; EGF-like.
CC InterPro: IPR001841; EGF_Ca.
CC InterPro: IPR001438; EGF_II.
CC InterPro: IPR002212; TB.
CC InterPro: IPR000822; Znf-C2H2.
CC Pfam: PF00008; EGF; 46.
CC Pfam: PF00683; TB; 9.
CC PRINTS: PR00010; EGFBLD.

```

DR	SMART: SMO0179; EGF_CA; 43.
DR	SMART: SMO0001; EGF_Like; 3.
DR	PROSITE: PS00010; ASX_HYDROXYL; 43.
DR	PROSITE: PS00022; EGF_1; 2.
DR	PROSITE: PS0186; POG_2; 36.
DR	PROSITE: PS0187; EGF_CA; 43.
KW	Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
KV	Repeat: Signal; Multigene family.
FT	SIGNAL 1 28 POTENTIAL 2.
FT	CHAIN 29 297 FIBRILLIN 2.
FT	DOMAIN 111 142 EGF-LIKE 1, NON-CALCIUM BINDING.
FT	DOMAIN 145 176 EGF-LIKE 2, NON-CALCIUM BINDING.
FT	DOMAIN 176 208 EGF-LIKE 3, NON-CALCIUM BINDING.
FT	DOMAIN 276 317 EGF-LIKE 4, CALCIUM-BINDING.
FT	DOMAIN 318 359 EGF-LIKE 5, CALCIUM-BINDING.
FT	REPEAT 360 426 TGFBP 1.
FT	DOMAIN 487 527 EGF-LIKE 6, NON-CALCIUM BINDING.
FT	DOMAIN 528 567 EGF-LIKE 7, CALCIUM-BINDING.
FT	DOMAIN 568 609 EGF-LIKE 8, CALCIUM-BINDING.
FT	DOMAIN 610 650 EGF-LIKE 9, CALCIUM-BINDING.
FT	DOMAIN 651 691 EGF-LIKE 10, CALCIUM-BINDING.
FT	REPEAT 692 760 TGFBP 2.
FT	DOMAIN 761 802 EGF-LIKE 11, CALCIUM-BINDING.
FT	DOMAIN 803 844 EGF-LIKE 12, CALCIUM-BINDING.
FT	DOMAIN 845 883 EGF-LIKE 13, CALCIUM-BINDING.
FT	DOMAIN 948 989 EGF-LIKE 14, CALCIUM-BINDING.
FT	REPEAT 990 1065 TGFBP 3.
FT	DOMAIN 1066 1107 EGF-LIKE 15, CALCIUM-BINDING.
FT	DOMAIN 1108 1150 EGF-LIKE 16, CALCIUM-BINDING.
FT	DOMAIN 1151 1192 EGF-LIKE 17, CALCIUM-BINDING.
FT	DOMAIN 1193 1234 EGF-LIKE 18, CALCIUM-BINDING.
FT	DOMAIN 1235 1275 EGF-LIKE 19, CALCIUM-BINDING.
FT	DOMAIN 1276 1317 EGF-LIKE 20, CALCIUM-BINDING.
FT	DOMAIN 1318 1359 EGF-LIKE 21, CALCIUM-BINDING.
FT	DOMAIN 1360 1400 EGF-LIKE 22, CALCIUM-BINDING.
FT	DOMAIN 1401 1441 EGF-LIKE 23, CALCIUM-BINDING.
FT	DOMAIN 1442 1483 EGF-LIKE 24, CALCIUM-BINDING.
FT	DOMAIN 1484 1524 EGF-LIKE 25, CALCIUM-BINDING.
FT	DOMAIN 1525 1565 EGF-LIKE 26, CALCIUM-BINDING.
FT	REPEAT 1566 1642 TGFBP 4.
FT	DOMAIN 1643 1684 EGF-LIKE 27, CALCIUM-BINDING.
FT	DOMAIN 1685 1726 EGF-LIKE 28, CALCIUM-BINDING.
FT	REPEAT 1727 1800 TGFBP 5.
FT	DOMAIN 1801 1842 EGF-LIKE 29, CALCIUM-BINDING.
FT	DOMAIN 1843 1884 EGF-LIKE 30, CALCIUM-BINDING.
FT	DOMAIN 1885 1926 EGF-LIKE 31, CALCIUM-BINDING.
FT	DOMAIN 1927 1965 EGF-LIKE 32, CALCIUM-BINDING.
FT	DOMAIN 1966 2008 EGF-LIKE 33, CALCIUM-BINDING.
FT	DOMAIN 2009 2048 EGF-LIKE 34, CALCIUM-BINDING.
FT	DOMAIN 2049 2090 EGF-LIKE 35, CALCIUM-BINDING.
FT	REPEAT 2091 2163 TGFBP 6.
FT	DOMAIN 2164 2205 EGF-LIKE 36, CALCIUM-BINDING.
FT	DOMAIN 2206 2245 EGF-LIKE 37, CALCIUM-BINDING.
FT	DOMAIN 2246 2286 EGF-LIKE 38, CALCIUM-BINDING.
FT	DOMAIN 2287 2330 EGF-LIKE 39, CALCIUM-BINDING.
FT	DOMAIN 2331 2372 EGF-LIKE 40, CALCIUM-BINDING.
FT	REPEAT 2373 2443 TGFBP 7.
FT	DOMAIN 2442 2483 EGF-LIKE 41, CALCIUM-BINDING.
FT	DOMAIN 2484 2524 EGF-LIKE 42, CALCIUM-BINDING.
FT	DOMAIN 2525 2563 EGF-LIKE 43, CALCIUM-BINDING.
FT	DOMAIN 2564 2606 EGF-LIKE 44, CALCIUM-BINDING.
FT	DOMAIN 2607 2646 EGF-LIKE 45, CALCIUM-BINDING.
FT	DOMAIN 2647 2687 EGF-LIKE 46, CALCIUM-BINDING.
FT	DOMAIN 2688 2727 EGF-LIKE 47, CALCIUM-BINDING.
FT	DISULFID 115 124 BY SIMILARITY.
FT	DISULFID 119 130 BY SIMILARITY.
FT	DISULFID 132 141 BY SIMILARITY.
FT	DISULFID 149 159 BY SIMILARITY.
FT	DISULFID 153 164 BY SIMILARITY.
FT	DISULFID 166 175 BY SIMILARITY.
FT	DISULFID 180 190 BY SIMILARITY.
FT	DISULFID 184 196 BY SIMILARITY.
FT	DISULFID 198 207 BY SIMILARITY.

FT	DISLFLTD	280	292	BY	SIMILARTTY
FT	DISLFLTD	287	301	BY	SIMILARTTY
FT	DISLFLTD	303	316	BY	SIMILARTTY
FT	DISLFLTD	322	334	BY	SIMILARTTY
FT	DISLFLTD	329	343	BY	SIMILARTTY
FT	DISLFLTD	345	358	BY	SIMILARTTY
FT	DISLFLTD	491	503	BY	SIMILARTTY
FT	DISLFLTD	498	512	BY	SIMILARTTY
FT	DISLFLTD	514	526	BY	SIMILARTTY
FT	DISLFLTD	532	542	BY	SIMILARTTY
FT	DISLFLTD	537	551	BY	SIMILARTTY
FT	DISLFLTD	553	566	BY	SIMILARTTY
FT	DISLFLTD	572	584	BY	SIMILARTTY
FT	DISLFLTD	579	593	BY	SIMILARTTY
FT	DISLFLTD	595	608	BY	SIMILARTTY
FT	DISLFLTD	614	625	BY	SIMILARTTY
FT	DISLFLTD	620	634	BY	SIMILARTTY
FT	DISLFLTD	636	649	BY	SIMILARTTY
FT	DISLFLTD	655	666	BY	SIMILARTTY
FT	DISLFLTD	661	675	BY	SIMILARTTY
FT	DISLFLTD	677	690	BY	SIMILARTTY
FT	DISLFLTD	765	777	BY	SIMILARTTY
FT	DISLFLTD	772	786	BY	SIMILARTTY
FT	DISLFLTD	788	801	BY	SIMILARTTY
FT	DISLFLTD	807	819	BY	SIMILARTTY
FT	DISLFLTD	814	828	BY	SIMILARTTY
FT	DISLFLTD	830	843	BY	SIMILARTTY
FT	DISLFLTD	849	859	BY	SIMILARTTY
FT	DISLFLTD	854	868	BY	SIMILARTTY
FT	DISLFLTD	870	883	BY	SIMILARTTY
FT	DISLFLTD	922	964	BY	SIMILARTTY
FT	DISLFLTD	959	973	BY	SIMILARTTY
FT	DISLFLTD	975	988	BY	SIMILARTTY
FT	DISLFLTD	1070	1082	BY	SIMILARTTY
FT	DISLFLTD	1077	1091	BY	SIMILARTTY
FT	DISLFLTD	1093	1106	BY	SIMILARTTY
FT	DISLFLTD	1112	1124	BY	SIMILARTTY
FT	DISLFLTD	1119	1133	BY	SIMILARTTY
FT	DISLFLTD	1135	1149	BY	SIMILARTTY
FT	DISLFLTD	1155	1167	BY	SIMILARTTY
FT	DISLFLTD	1162	1176	BY	SIMILARTTY
FT	DISLFLTD	1178	1191	BY	SIMILARTTY
FT	DISLFLTD	1197	1209	BY	SIMILARTTY
FT	DISLFLTD	1204	1218	BY	SIMILARTTY
FT	DISLFLTD	1220	1233	BY	SIMILARTTY
FT	DISLFLTD	1239	1250	BY	SIMILARTTY
FT	DISLFLTD	1246	1259	BY	SIMILARTTY
FT	DISLFLTD	1261	1274	BY	SIMILARTTY
FT	DISLFLTD	1280	1292	BY	SIMILARTTY
FT	DISLFLTD	1287	1301	BY	SIMILARTTY
FT	DISLFLTD	1303	1316	BY	SIMILARTTY
FT	DISLFLTD	1322	1334	BY	SIMILARTTY
FT	DISLFLTD	1339	1343	BY	SIMILARTTY
FT	DISLFLTD	1345	1358	BY	SIMILARTTY
FT	DISLFLTD	1364	1377	BY	SIMILARTTY
FT	DISLFLTD	1371	1386	BY	SIMILARTTY
FT	DISLFLTD	1388	1399	BY	SIMILARTTY
FT	DISLFLTD	1405	1418	BY	SIMILARTTY
FT	DISLFLTD	1412	1427	BY	SIMILARTTY
FT	DISLFLTD	1429	1440	BY	SIMILARTTY
FT	DISLFLTD	1446	1458	BY	SIMILARTTY
FT	DISLFLTD	1453	1467	BY	SIMILARTTY
FT	DISLFLTD	1469	1482	BY	SIMILARTTY
FT	DISLFLTD	1488	1499	BY	SIMILARTTY
FT	DISLFLTD	1494	1508	BY	SIMILARTTY
FT	DISLFLTD	1510	1523	BY	SIMILARTTY
FT	DISLFLTD	1529	1540	BY	SIMILARTTY
FT	DISLFLTD	1535	1549	BY	SIMILARTTY
FT	DISLFLTD	1551	1564	BY	SIMILARTTY
FT	DISLFLTD	1547	1659	BY	SIMILARTTY
FT	DISLFLTD	1644	1668	BY	SIMILARTTY
FT	DISLFLTD	1670	1683	BY	SIMILARTTY
FT	DISLFLTD	1689	1701	BY	SIMILARTTY


```

FT DOMAIN 1241 1281 EGF-LIKE 19, CALCIUM-BINDING.
FT DOMAIN 1282 1323 EGF-LIKE 20, CALCIUM-BINDING.
FT DOMAIN 1324 1365 EGF-LIKE 21, CALCIUM-BINDING.
FT DOMAIN 1366 1406 EGF-LIKE 22, CALCIUM-BINDING.
FT DOMAIN 1407 1447 EGF-LIKE 23, CALCIUM-BINDING.
FT DOMAIN 1448 1489 EGF-LIKE 24, CALCIUM-BINDING.
FT DOMAIN 1490 1530 EGF-LIKE 25, CALCIUM-BINDING.
FT DOMAIN 1531 1571 EGF-LIKE 26, CALCIUM-BINDING.
FT REPEAT 1572 1648 TGFBRP 4.
FT DOMAIN 1649 1690 EGF-LIKE 27, CALCIUM-BINDING.
FT DOMAIN 1691 1732 EGF-LIKE 28, CALCIUM-BINDING.
FT REPEAT 1733 1806 TGFBRP 5.
FT DOMAIN 1807 1848 EGF-LIKE 29, CALCIUM-BINDING.
FT DOMAIN 1849 1890 EGF-LIKE 30, CALCIUM-BINDING.
FT DOMAIN 1891 1932 EGF-LIKE 31, CALCIUM-BINDING.
FT DOMAIN 1933 1971 EGF-LIKE 32, CALCIUM-BINDING.
FT DOMAIN 1972 2014 EGF-LIKE 33, CALCIUM-BINDING.
FT DOMAIN 2015 2054 EGF-LIKE 34, CALCIUM-BINDING.
FT DOMAIN 2055 2096 EGF-LIKE 35, CALCIUM-BINDING.
FT REPEAT 2097 2169 TGFBRP 6.
FT DOMAIN 2170 2211 EGF-LIKE 36, CALCIUM-BINDING.
FT DOMAIN 2212 2251 EGF-LIKE 37, CALCIUM-BINDING.
FT DOMAIN 2252 2292 EGF-LIKE 38, CALCIUM-BINDING.
FT DOMAIN 2293 2336 EGF-LIKE 39, CALCIUM-BINDING.
FT DOMAIN 2337 2378 EGF-LIKE 40, CALCIUM-BINDING.
FT REPEAT 2379 2447 TGFBRP 7.
FT DOMAIN 2448 2489 EGF-LIKE 41, CALCIUM-BINDING.
FT DOMAIN 2490 2530 EGF-LIKE 42, CALCIUM-BINDING.
FT DOMAIN 2531 2569 EGF-LIKE 43, CALCIUM-BINDING.
FT DOMAIN 2570 2612 EGF-LIKE 44, CALCIUM-BINDING.
FT DOMAIN 2613 2652 EGF-LIKE 45, CALCIUM-BINDING.
FT DOMAIN 2653 2693 EGF-LIKE 46, CALCIUM-BINDING.
FT DOMAIN 2694 2733 EGF-LIKE 47, CALCIUM-BINDING.
FT DOMAIN 2734 2773 EGF-LIKE 48, CALCIUM-BINDING.
FT DOMAIN 2774 2813 EGF-LIKE 49, CALCIUM-BINDING.
FT DOMAIN 2814 2853 EGF-LIKE 50, CALCIUM-BINDING.
FT DOMAIN 2854 2893 EGF-LIKE 51, CALCIUM-BINDING.
FT DOMAIN 2894 2933 EGF-LIKE 52, CALCIUM-BINDING.
FT DOMAIN 2934 2973 EGF-LIKE 53, CALCIUM-BINDING.
FT DOMAIN 2974 3013 EGF-LIKE 54, CALCIUM-BINDING.
FT DOMAIN 3014 3053 EGF-LIKE 55, CALCIUM-BINDING.
FT DOMAIN 3054 3093 EGF-LIKE 56, CALCIUM-BINDING.
FT DOMAIN 3094 3133 EGF-LIKE 57, CALCIUM-BINDING.
FT DOMAIN 3134 3173 EGF-LIKE 58, CALCIUM-BINDING.
FT DOMAIN 3174 3213 EGF-LIKE 59, CALCIUM-BINDING.
FT DOMAIN 3214 3253 EGF-LIKE 60, CALCIUM-BINDING.
FT DOMAIN 3254 3293 EGF-LIKE 61, CALCIUM-BINDING.
FT DOMAIN 3294 3333 EGF-LIKE 62, CALCIUM-BINDING.
FT DOMAIN 3334 3373 EGF-LIKE 63, CALCIUM-BINDING.
FT DOMAIN 3374 3413 EGF-LIKE 64, CALCIUM-BINDING.
FT DOMAIN 3414 3453 EGF-LIKE 65, CALCIUM-BINDING.
FT DOMAIN 3454 3493 EGF-LIKE 66, CALCIUM-BINDING.
FT DOMAIN 3494 3533 EGF-LIKE 67, CALCIUM-BINDING.
FT DOMAIN 3534 3573 EGF-LIKE 68, CALCIUM-BINDING.
FT DOMAIN 3574 3613 EGF-LIKE 69, CALCIUM-BINDING.
FT DOMAIN 3614 3653 EGF-LIKE 70, CALCIUM-BINDING.
FT DOMAIN 3654 3693 EGF-LIKE 71, CALCIUM-BINDING.
FT DOMAIN 3694 3733 EGF-LIKE 72, CALCIUM-BINDING.
FT DOMAIN 3734 3773 EGF-LIKE 73, CALCIUM-BINDING.
FT DOMAIN 3774 3813 EGF-LIKE 74, CALCIUM-BINDING.
FT DOMAIN 3814 3853 EGF-LIKE 75, CALCIUM-BINDING.
FT DOMAIN 3854 3893 EGF-LIKE 76, CALCIUM-BINDING.
FT DOMAIN 3894 3933 EGF-LIKE 77, CALCIUM-BINDING.
FT DOMAIN 3934 3973 EGF-LIKE 78, CALCIUM-BINDING.
FT DOMAIN 3974 4013 EGF-LIKE 79, CALCIUM-BINDING.
FT DOMAIN 4014 4053 EGF-LIKE 80, CALCIUM-BINDING.
FT DOMAIN 4054 4093 EGF-LIKE 81, CALCIUM-BINDING.
FT DOMAIN 4094 4133 EGF-LIKE 82, CALCIUM-BINDING.
FT DOMAIN 4134 4173 EGF-LIKE 83, CALCIUM-BINDING.
FT DOMAIN 4174 4213 EGF-LIKE 84, CALCIUM-BINDING.
FT DOMAIN 4214 4253 EGF-LIKE 85, CALCIUM-BINDING.
FT DOMAIN 4254 4293 EGF-LIKE 86, CALCIUM-BINDING.
FT DOMAIN 4294 4333 EGF-LIKE 87, CALCIUM-BINDING.
FT DOMAIN 4334 4373 EGF-LIKE 88, CALCIUM-BINDING.
FT DOMAIN 4374 4413 EGF-LIKE 89, CALCIUM-BINDING.
FT DOMAIN 4414 4453 EGF-LIKE 90, CALCIUM-BINDING.

```

```

FT DISULFID 965 979 BY SIMILARITY.
FT DISULFID 981 994 BY SIMILARITY.
FT DISULFID 1076 1088 BY SIMILARITY.
FT DISULFID 1083 1097 BY SIMILARITY.
FT DISULFID 1099 1112 BY SIMILARITY.
FT DISULFID 1118 1130 BY SIMILARITY.
FT DISULFID 1141 1153 BY SIMILARITY.
FT DISULFID 1145 1155 BY SIMILARITY.
FT DISULFID 1161 1173 BY SIMILARITY.
FT DISULFID 1168 1182 BY SIMILARITY.
FT DISULFID 1184 1197 BY SIMILARITY.
FT DISULFID 1203 1215 BY SIMILARITY.
FT DISULFID 1210 1224 BY SIMILARITY.

Query Match 13.4% Score 409.5; DB 1; Length 2911;
Best Local Similarity 28.6%; Pred. No. 7.6e-21;
Matches 96; Conservative 37; Mismatches 92; Indels 111; Gaps 13;

QY 38 PGVCHYGT-----KLACCYGW----RRNSKGVCEATGEPG-CRFGCV--GPNKC 80
DB 1080 PGMCTYGCRCNTIGSFKCRNSGFALDMEERNCTDIDECRISPDLCGSGICVNTPGSFEC 1139
QY 81 RCFPEY-----TGKTSQ-----D 94
DB 1140 ECFEBSGFFMMKNCMDIDGECERNPLLCRGCTCVNTEGSPQCCPLGHELSPSRDCVD 1199
QY 95 VNECGMKRPPCOH-RCVNTGHSYKFCFLSGHMLMPDAT-CVNSRTCAMIN--COYSCEDT 150
DB 1200 INECSLSNLCRNGCVNMIGTYQCSNPYQAPRDRGCTDIDECMIMNGSCDPTQCTNS 1259
QY 151 EEGPQCLPSSGRLRAPNGRCLDIDECASGRVY-----PYNRGV-----192
DB 1260 EGSYECSC-SBGYALMPDRGSCADIDECENNPDIODGQCTNIPGEYRCLYDGFMSMD 1318
QY 193 -----NTFGSYCKCHIGFELQYISGRDIDINECTMDSHT 229
DB 1319 MKTCIDVMECDLNSNICMFGECENTKGSFICHQCLGYVKK--KGTGTVDYDEEIGHN 1376
QY 230 CSHANCFNTGSGFKCKCKGKNGSLRCSAIPENS 265
DB 1377 CDMHASCINIGSFKSCREGMIGIKICIDIDEC 1412

RESULT 3
FBN1_BOVIN STANDARD; PRT; 2871 AA.
ID FBN1_BOVIN
AC P98133;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fibrillin 1 precursor (MP340).
GN FBN1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCB1_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=95137597; PubMed=7835900;
RA Tilstra D.J., Potter K.A., Byers P.H.;
RT "Sequence of the coding region of the bovine fibrillin cDNA and
RT localization to bovine chromosome 10.";
RL Genomics 23:480-485(1994).
RP PARTIAL SEQUENCE.
RX MEDLINE=96132851; PubMed=8557636;
RA Gibson M.A., Hatzilakolas G., Kumaratilake J.S., Sandberg L.B.,
RA Nicholl J.K., Sutherland G.R., Cleary E.G.;
RT "Further characterization of proteins associated with elastic fiber
RT microfibrils including the molecular cloning of MAGP-2 (MP25).";
RL J. Biol. Chem. 271:1096-1103(1996).

```


FT DISULFID 1182 1195 BY SIMILARITY.
 FT DISULFID 1201 1212 BY SIMILARITY.
 FT DISULFID 1208 1221 BY SIMILARITY.
 FT DISULFID 1223 1236 BY SIMILARITY.
 FT DISULFID 1242 1254 BY SIMILARITY.
 FT DISULFID 1249 1263 BY SIMILARITY.
 FT DISULFID 1265 1278 BY SIMILARITY.
 FT DISULFID 1284 1296 BY SIMILARITY.
 FT DISULFID 1291 1305 BY SIMILARITY.
 FT DISULFID 1307 1320 BY SIMILARITY.
 FT DISULFID 1326 1339 BY SIMILARITY.
 FT DISULFID 1333 1348 BY SIMILARITY.
 FT DISULFID 1350 1361 BY SIMILARITY.
 FT DISULFID 1367 1380 BY SIMILARITY.
 FT DISULFID 1374 1389 BY SIMILARITY.
 FT DISULFID 1391 1402 BY SIMILARITY.
 FT DISULFID 1408 1420 BY SIMILARITY.
 FT DISULFID 1415 1429 BY SIMILARITY.
 FT DISULFID 1431 1444 BY SIMILARITY.
 FT DISULFID 1450 1461 BY SIMILARITY.
 FT DISULFID 1456 1470 BY SIMILARITY.
 FT DISULFID 1472 1485 BY SIMILARITY.
 FT DISULFID 1491 1502 BY SIMILARITY.
 FT DISULFID 1497 1511 BY SIMILARITY.
 FT DISULFID 1513 1526 BY SIMILARITY.
 FT DISULFID 1610 1622 BY SIMILARITY.

Query Match 12.7% Score 387.5; DB 1; Length 2871;
 Best Local Similarity 30.6%; Pred. No. 2.6e-19;
 Matches 89; Conservative 34; Mismatches 91; Indels 77; Gaps 10;

QY 49 CCYGRNRKSKVCATCPG-----CAFGCV--GPNKRC 82
 DB 1081 CGRQCVPNTPDPEFCCKCEGESGFMNKNMCDIDECORDELKCGVCLMTGEGSYREC 1140
 QY 83 PPGY---TGKCSODVNEGCMKPRPCOH-RCVNTGSKYKFCCLSGHMLMPAT-CVNSRT 137
 DB 1141 PPGHDLANINACIDINCELSAHLCPHRCVNLGKQCACNPGYHSTPRLFCVDDIDE 1200
 QY 138 CAMIN--CQYSCDEPTBEGPQCLPSSGLRLAPNGRDCLDIDECASGVKIC-----PY 187
 DB 1201 CSIMNGGCTFCTJNSEGSEYSCSC-OPGFALMPDORSCSTDIDECEDNPNICDGGGCTNIRG 1259
 QY 188 NRRCV-----NTGSGYCKKHIGFELQYISGR 214
 DB 1260 EYRLCYDGFMASEDMKTCVDVNECDLNPNICLSGTCENTGSGFICHCMDGYSGR--RGK 1317
 QY 215 YDCIDINECTDMSHTCSHANCFTPGSFCKCKGQYKNGLRCSAIPENS 265
 DB 1318 TGCIDINCEIGAHNCDRAVCTNAGSFCKSCSPGWIIGDGIKTDLECS 1368

RESULT 4
 FBNI_MOUSE STANDARD; PRT; 2871 AA.
 AC Q61554; 060826;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Fibrillin 1 precursor.
 GN FBNI OR FBNI-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN 1;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95130561; PubMed=7829516;
 RA Yin W., Germiller J., Sanguinetti C., Smiley E., Pangillian T.,
 RA Pereira L., Ramirez F., Bonadio J.;
 RT "Primary structure and developmental expression of Fbn-1, the mouse
 fibrillin gene";
 RL J. Biol. Chem. 270:11798-1806(1995).

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CD-1; TISSUE=Kidney;
 RA Ota K., Kumar A., Wada J.; Kanwar Y.S.;
 RL Submitted (APR-1995) to the EMBL/GenBank/DDBJ databases.
 CC -1- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
 CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE
 CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
 CC -1- PTM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER
 CC FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE
 CC MICROFIBRILS (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 43 CALCIUM-BINDING AND 4 NON-CALCIUM BINDING
 CC EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 DR EMBL: L29454; AAA56840.1; -;
 DR EMBL: U22493; AAA64217.1; -;
 DR HSSP: P35353; IAPJ.
 DR MGD: MGI:95489; Fbn1.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF-Ca.
 DR InterPro: IPR002212; TB.
 DR Pfam: PF00006; EGF; 46.
 DR Pfam: PF00683; TB; 9.
 DR SMART: SM00179; EGF_CA; 42.
 DR SMART: SM00001; EGF_1like; 4.
 DR PROSITE: PS00010; ASX_HYDROXYL; 43.
 DR PROSITE: PS00022; EGF_1; 2.
 DR PROSITE: PS01186; EGF_2; 38.
 DR PROSITE: PS01187; EGF_CA; 45.
 KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
 KW Repeat; signal; Multigene family.
 FT SIGNAL 1 27
 FT CHAIN 28 2871
 FT DOMAIN 81 112 FIBRILLIN 1.
 FT DOMAIN 115 146 EGF-LIKE 1, NON-CALCIUM BINDING.
 FT DOMAIN 147 178 EGF-LIKE 2, NON-CALCIUM BINDING.
 FT DOMAIN 246 287 EGF-LIKE 3, NON-CALCIUM BINDING.
 FT DOMAIN 288 329 EGF-LIKE 4, CALCIUM-BINDING.
 FT DOMAIN 330 401 EGF-LIKE 5, CALCIUM-BINDING.
 FT DOMAIN 402 446 TGFBRP 1.
 FT DOMAIN 449 489 PRO-RICH.
 FT DOMAIN 490 529 EGF-LIKE 6, NON-CALCIUM BINDING.
 FT DOMAIN 530 571 EGF-LIKE 7, CALCIUM-BINDING.
 FT DOMAIN 572 612 EGF-LIKE 8, CALCIUM-BINDING.
 FT DOMAIN 613 653 EGF-LIKE 9, CALCIUM-BINDING.
 FT DOMAIN 656 721 EGF-LIKE 10, CALCIUM-BINDING.
 FT DOMAIN 723 764 TGFBRP 2.
 FT DOMAIN 765 806 EGF-LIKE 11, CALCIUM-BINDING.
 FT DOMAIN 807 846 EGF-LIKE 12, CALCIUM-BINDING.
 FT DOMAIN 910 951 EGF-LIKE 13, CALCIUM-BINDING.
 FT DOMAIN 952 1018 EGF-LIKE 14, CALCIUM-BINDING.
 FT DOMAIN 1028 1069 TGFBRP 3.
 FT DOMAIN 1070 1112 EGF-LIKE 15, CALCIUM-BINDING.
 FT DOMAIN 1113 1154 EGF-LIKE 16, CALCIUM-BINDING.
 FT DOMAIN 1155 1196 EGF-LIKE 17, CALCIUM-BINDING.
 FT DOMAIN 1197 1237 EGF-LIKE 18, CALCIUM-BINDING.
 FT DOMAIN 1238 1279 EGF-LIKE 19, CALCIUM-BINDING.
 FT DOMAIN 1280 1321 EGF-LIKE 20, CALCIUM-BINDING.
 FT DOMAIN 1321 1362 EGF-LIKE 21, CALCIUM-BINDING.
 FT DOMAIN 1362 1403 EGF-LIKE 22, CALCIUM-BINDING.
 FT DOMAIN 1403 1445 EGF-LIKE 23, CALCIUM-BINDING.
 FT DOMAIN 1445 1486 EGF-LIKE 24, CALCIUM-BINDING.
 FT DOMAIN 1487 1527 EGF-LIKE 25, CALCIUM-BINDING.
 FT DOMAIN 1527 EGF-LIKE 26, CALCIUM-BINDING.

FT	DOMAIN	1528	1599	TGFBP 4.
FT	DOMAIN	1606	1647	EGF-LIKE 27, CALCIUM-BINDING.
FT	DOMAIN	1648	1688	EGF-LIKE 28, CALCIUM-BINDING.
FT	DOMAIN	1689	1758	TGFBP 5.
FT	DOMAIN	1766	1807	EGF-LIKE 29, CALCIUM-BINDING.
FT	DOMAIN	1808	1848	EGF-LIKE 30, CALCIUM-BINDING.
FT	DOMAIN	1849	1890	EGF-LIKE 31, CALCIUM-BINDING.
FT	DOMAIN	1891	1929	EGF-LIKE 32, CALCIUM-BINDING.
FT	DOMAIN	1930	1972	EGF-LIKE 33, CALCIUM-BINDING.
FT	DOMAIN	1973	2012	EGF-LIKE 34, CALCIUM-BINDING.
FT	DOMAIN	2013	2034	EGF-LIKE 35, CALCIUM-BINDING.
FT	DOMAIN	2055	2121	TGFBP 6.
FT	DOMAIN	2127	2165	EGF-LIKE 36, CALCIUM-BINDING.
FT	DOMAIN	2166	2205	EGF-LIKE 37, CALCIUM-BINDING.
FT	DOMAIN	2206	2246	EGF-LIKE 38, CALCIUM-BINDING.
FT	DOMAIN	2247	2290	EGF-LIKE 39, CALCIUM-BINDING.
FT	DOMAIN	2291	2332	EGF-LIKE 40, CALCIUM-BINDING.
FT	DOMAIN	2333	2400	TGFBP 7.
FT	DOMAIN	2402	2443	EGF-LIKE 41, CALCIUM-BINDING.
FT	DOMAIN	2444	2484	EGF-LIKE 42, CALCIUM-BINDING.
FT	DOMAIN	2485	2523	EGF-LIKE 43, CALCIUM-BINDING.
FT	DOMAIN	2524	2566	EGF-LIKE 44, CALCIUM-BINDING.
FT	DOMAIN	2567	2606	EGF-LIKE 45, CALCIUM-BINDING.
FT	DOMAIN	2607	2647	EGF-LIKE 46, CALCIUM-BINDING.
FT	DOMAIN	2648	2687	EGF-LIKE 47, CALCIUM-BINDING.
FT	DISULFID	85	94	BY SIMILARITY.
FT	DISULFID	89	100	BY SIMILARITY.
FT	DISULFID	102	111	BY SIMILARITY.
FT	DISULFID	119	129	BY SIMILARITY.
FT	DISULFID	123	134	BY SIMILARITY.
FT	DISULFID	136	145	BY SIMILARITY.
FT	DISULFID	150	160	BY SIMILARITY.
FT	DISULFID	154	166	BY SIMILARITY.
FT	DISULFID	168	177	BY SIMILARITY.
FT	DISULFID	250	262	BY SIMILARITY.
FT	DISULFID	257	271	BY SIMILARITY.
FT	DISULFID	273	286	BY SIMILARITY.
FT	DISULFID	292	304	BY SIMILARITY.
FT	DISULFID	299	313	BY SIMILARITY.
FT	DISULFID	315	328	BY SIMILARITY.
FT	DISULFID	453	465	BY SIMILARITY.
FT	DISULFID	460	474	BY SIMILARITY.
FT	DISULFID	476	488	BY SIMILARITY.
FT	DISULFID	494	504	BY SIMILARITY.
FT	DISULFID	499	513	BY SIMILARITY.
FT	DISULFID	515	528	BY SIMILARITY.
FT	DISULFID	534	546	BY SIMILARITY.
FT	DISULFID	541	555	BY SIMILARITY.
FT	DISULFID	557	570	BY SIMILARITY.
FT	DISULFID	576	587	BY SIMILARITY.
FT	DISULFID	582	596	BY SIMILARITY.
FT	DISULFID	598	611	BY SIMILARITY.
FT	DISULFID	617	628	BY SIMILARITY.
FT	DISULFID	623	637	BY SIMILARITY.
FT	DISULFID	639	652	BY SIMILARITY.
FT	DISULFID	727	739	BY SIMILARITY.
FT	DISULFID	734	748	BY SIMILARITY.
FT	DISULFID	750	763	BY SIMILARITY.
FT	DISULFID	769	781	BY SIMILARITY.
FT	DISULFID	776	790	BY SIMILARITY.
FT	DISULFID	792	805	BY SIMILARITY.
FT	DISULFID	811	821	BY SIMILARITY.
FT	DISULFID	816	830	BY SIMILARITY.
FT	DISULFID	832	845	BY SIMILARITY.
FT	DISULFID	914	926	BY SIMILARITY.
FT	DISULFID	921	935	BY SIMILARITY.
FT	DISULFID	937	950	BY SIMILARITY.
FT	DISULFID	1032	1044	BY SIMILARITY.
FT	DISULFID	1039	1053	BY SIMILARITY.
FT	DISULFID	1055	1068	BY SIMILARITY.
FT	DISULFID	1074	1086	BY SIMILARITY.
FT	DISULFID	1081	1095	BY SIMILARITY.
FT	DISULFID	1097	1111	BY SIMILARITY.

FT	DISULFID	1117	1129	BY SIMILARITY.
FT	DISULFID	1124	1138	BY SIMILARITY.
FT	DISULFID	1140	1153	BY SIMILARITY.
FT	DISULFID	1159	1171	BY SIMILARITY.
FT	DISULFID	1166	1180	BY SIMILARITY.
FT	DISULFID	1182	1195	BY SIMILARITY.
FT	DISULFID	1201	1212	BY SIMILARITY.
FT	DISULFID	1208	1221	BY SIMILARITY.
FT	DISULFID	1223	1236	BY SIMILARITY.
FT	DISULFID	1242	1254	BY SIMILARITY.
FT	DISULFID	1249	1263	BY SIMILARITY.
FT	DISULFID	1265	1278	BY SIMILARITY.
FT	DISULFID	1284	1296	BY SIMILARITY.
FT	DISULFID	1301	1305	BY SIMILARITY.
FT	DISULFID	1307	1320	BY SIMILARITY.
FT	DISULFID	1326	1339	BY SIMILARITY.
FT	DISULFID	1333	1348	BY SIMILARITY.
FT	DISULFID	1350	1361	BY SIMILARITY.
FT	DISULFID	1367	1380	BY SIMILARITY.
FT	DISULFID	1374	1389	BY SIMILARITY.
FT	DISULFID	1391	1402	BY SIMILARITY.
FT	DISULFID	1408	1420	BY SIMILARITY.
FT	DISULFID	1415	1429	BY SIMILARITY.
FT	DISULFID	1431	1444	BY SIMILARITY.
FT	DISULFID	1450	1461	BY SIMILARITY.
FT	DISULFID	1456	1470	BY SIMILARITY.
FT	DISULFID	1472	1485	BY SIMILARITY.
FT	DISULFID	1491	1502	BY SIMILARITY.
FT	DISULFID	1497	1511	BY SIMILARITY.
FT	DISULFID	1513	1526	BY SIMILARITY.
FT	DISULFID	1610	1622	BY SIMILARITY.
FT	DISULFID	1617	1631	BY SIMILARITY.
FT	DISULFID	1633	1646	BY SIMILARITY.
FT	DISULFID	1652	1663	BY SIMILARITY.
FT	DISULFID	1658	1672	BY SIMILARITY.

Query Match 12.6% Score 385.5; DB 1; Length 2871;
 Best Local Similarity 32.2%, Pred. No. 3,7e-19;
 Matches 86; Conservative 34; Mismatches 82; Indels 65; Gaps 11;

QY	39	GVCHYGTKLACCYGMRNSKGVCEATCEPG-----CKFGECV-75
DB	1127	GICH-----TTEGTYRCECPGHQLSPNISACIDINECELSANICPGRVCVN-1173
QY	76	--GPNKCKCPFGY---TCKTCSODVNECGMKRPPQOHNCVMTGHSYKGFCLSGHMLMDA-130
DB	1174	LIGKYQACNPGYHPTHDRLFCVDDIECSIMNGCETNSDGSYECSCOPGFALMPD-1232
QY	131	TCVNSRTFCAMINCOYSCEDTE---EGPQ-----CLCPSSGRLAPNGRDCIDIDEC-178
DB	1233	---QRSCSTDID---QCEDNINICDGGCCTNIPGEYRLC-YDGFMASEDMKTCVDVNEC-1284
QY	179	ASGVICYPNRCVNTFSGSYCKCHIGFELYISGRYCIDINECTMDSHYCSHHANCFN-238
DB	1285	DLPNPTIC-LSGTCENTKCSFICHCDMGYSGR--KGKTCFTDINECEIGHNCGRAVCTN-1341
QY	239	TGSPFKCKXQGYKNGLRCSAIPENS-265
DB	1342	TAGSFKCSCSPGWIGDGIKCTIDLECS-1368

RESULT 5
 FBNI_PIG STANDARD; PRT: 2871 AA.
 AC Q9TV36;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Fibrillin 1 precursor.
 GN FBNI.
 OS Sus scrofa (pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.


```

FT DISULFID 1039 1053 BY SIMILARITY.
FT DISULFID 1055 1068 BY SIMILARITY.
FT DISULFID 1074 1086 BY SIMILARITY.
FT DISULFID 1081 1095 BY SIMILARITY.
FT DISULFID 1097 1111 BY SIMILARITY.
FT DISULFID 1117 1129 BY SIMILARITY.
FT DISULFID 1124 1138 BY SIMILARITY.
FT DISULFID 1140 1153 BY SIMILARITY.
FT DISULFID 1159 1171 BY SIMILARITY.
FT DISULFID 1201 1212 BY SIMILARITY.
FT DISULFID 1208 1221 BY SIMILARITY.
FT DISULFID 1223 1236 BY SIMILARITY.
FT DISULFID 1242 1254 BY SIMILARITY.
FT DISULFID 1249 1263 BY SIMILARITY.
FT DISULFID 1265 1278 BY SIMILARITY.
FT DISULFID 1284 1296 BY SIMILARITY.
FT DISULFID 1291 1305 BY SIMILARITY.
FT DISULFID 1307 1320 BY SIMILARITY.
FT DISULFID 1326 1339 BY SIMILARITY.
FT DISULFID 1333 1348 BY SIMILARITY.
FT DISULFID 1350 1361 BY SIMILARITY.
FT DISULFID 1367 1380 BY SIMILARITY.
FT DISULFID 1374 1389 BY SIMILARITY.
FT DISULFID 1391 1402 BY SIMILARITY.
FT DISULFID 1408 1420 BY SIMILARITY.
FT DISULFID 1415 1429 BY SIMILARITY.
FT DISULFID 1450 1461 BY SIMILARITY.
FT DISULFID 1456 1470 BY SIMILARITY.
FT DISULFID 1472 1485 BY SIMILARITY.
FT DISULFID 1491 1502 BY SIMILARITY.
FT DISULFID 1497 1511 BY SIMILARITY.
FT DISULFID 1513 1526 BY SIMILARITY.
FT DISULFID 1610 1622 BY SIMILARITY.
FT DISULFID 1617 1631 BY SIMILARITY.
FT DISULFID 1633 1646 BY SIMILARITY.
FT DISULFID 1652 1663 BY SIMILARITY.
FT DISULFID 1658 1672 BY SIMILARITY.
FT DISULFID 1674 1687 BY SIMILARITY.
FT DISULFID 1770 1782 BY SIMILARITY.
FT DISULFID 1777 1791 BY SIMILARITY.
FT DISULFID 1793 1806 BY SIMILARITY.
FT DISULFID 1812 1824 BY SIMILARITY.
FT DISULFID 1818 1833 BY SIMILARITY.
FT DISULFID 1835 1847 BY SIMILARITY.

```

Query Match 12.6%; Score 385.5; DB 1; Length 2871;
 Best Local Similarity 27.2%; Pred. No. 3.7e-19;
 Matches 91; Conservative 38; Mismatches 84; Indels 121; Gaps 11;

```

QY 49 CCYGMRRNSKVCENTCEPG-----CKFGECV--GPNKRCF 83
DB 1039 CTGKCKNTIGSFKRCRDSGFDLSEERNCIDIDECRISPDLCGRGCVNTPGDFECKD 1098
QY 84 PGTYG-----KTSQDVNCEGKMPQCQHR-CVNFHSGYCFCLSGHMLMDAT----- 131
DB 1099 EGESEGMAMKNC-MDIDECQRPDLRCGVCLNTEGSEYCEPCSGHMSFNISACIDIN 1157
QY 132 -----CVNSRQCAMIN--COYSCEDTEE 152
DB 1158 ECELSAHLCPHRCVNLIGYQRRARNPGYSTPRLFCVDIDECSTINMGCEFTCTNSEG 1217
QY 153 GPQICPSSGLRLAPNGRCDLIDECASGVYIC-----PYNRRCV----- 192
DB 1218 SYEESC-QPQFALMPDRGCTDIDECEDNPNICOGGCTNIPGFRCLCYDGFVASEDMK 1276
QY 193 -----NPFSGYCKCHIGFELYISGRYDCIDINECTMDSHFCS 231
DB 1277 TCVDVNCEDLNPNICLSGTEGTFGSCFICHCMDGYSK--KGKTCYDINECEIGAHCND 1334
QY 232 HHANCFNTGSPFKCKGKGYKGNGLRCSAIPENS 265
DB 1335 RHAVCITYAGSFNCSGSPGWIIGDICKTDLDECS 1368

```

```

RESULT 6
FBNI_HUMAN STANDARD; PRT: 2871 AA.
ID FBNI_HUMAN
P35555;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fibrillin 1 precursor.
GN FBNI OR FBN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN 1;
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=93372860; PubMed=8364578;
RA Pereira L.V., D'Alessio M., Ramirez F., Lynch J.R., Sykes B.,
RA Pangillan T., Bonadio J.;
RT "Genomic organization of the sequence coding for fibrillin, the
RT defective gene product in Marfan syndrome.";
RL Hum. Mol. Genet. 2:961-968(1993).
RN 12;
RP SEQUENCE OF 1-932 FROM N.A.
RC TISSUE=Placenta, and fibroblast;
RX MEDLINE=94010947; PubMed=7691719;
RA Corson G.M., Chalberg S.C., Dietz H.C., Charbonneau N.L., Sakai L.Y.;
RT "Fibrillin binds calcium and is coded by cDNAs that reveal a
RT multidomain structure and alternatively spliced exons at the 5'
RT end.";
RL Genomics 17:476-484(1993).
RN 13;
RP SEQUENCE OF 899-2871 FROM N.A.
RX MEDLINE=91304567; PubMed=1852207;
RA Maslen C.L., Corson G.M., Maddox B.K., Glanville R.W., Sakai L.Y.;
RT "Partial sequence of a candidate gene for the Marfan syndrome.";
RL Nature 352:334-337(1991).
RN 14;
RP SEQUENCE OF 813-1313 FROM N.A.
RX MEDLINE=91304567; PubMed=1852206;
RA Lee B., Godfrey M., Vitale E., Hori H., Mattel M.-G., Sarfarazi M.,
RA Tsipouras P., Ramirez F., Hollister D.W.;
RT "Linkage of Marfan syndrome and a phenotypically related disorder to
RT two different fibrillin genes.";
RL Nature 352:330-334(1991).
RN 15;
RP CHARACTERIZATION.
RX MEDLINE=91317849; PubMed=1860873;
RA Sakai L.Y., Keene D.R., Glanville R.W., Bachinger H.P.;
RT "Purification and partial characterization of fibrillin, a cysteine-
RT rich structural component of connective tissue microfibrils.";
RL J. Biol. Chem. 266:14763-14770(1991).
RN 16;
RP STRUCTURE BY NMR OF 2054-2125.
RX MEDLINE=98031893; PubMed=9362480;
RA Yuan X., Downing A.K., Knott V., Handford P.A.;
RT "Solution structure of the transforming growth factor beta-binding
RT protein-like module, a domain associated with matrix fibrils.";
RL EMBO J. 16:6659-6666(1997).
RN 17;
RP STRUCTURE BY NMR OF 2124-2205.
RX MEDLINE=96144829; PubMed=8568869;
RA Knott V., Downing A.K., Cardy C.M., Handford P.A.;
RT "Calcium binding properties of an epidermal growth factor-like domain
RT pair from human fibrillin-1.";
RL J. Mol. Biol. 255:22-27(1996).
RN 18;
RP STRUCTURE BY NMR OF 2124-2205.
RX MEDLINE=96222301; PubMed=8653794;
RA Downing A.K., Knott V., Werner J.M., Cardy C.M., Campbell I.D.,
RA Handford P.A.;
RT "Solution structure of a pair of calcium-binding epidermal growth

```


FT DISULFID 821 836 BY SIMILARITY.
 FT DISULFID 843 856 BY SIMILARITY.
 FT DISULFID 862 875 BY SIMILARITY.
 FT DISULFID 869 884 BY SIMILARITY.
 FT DISULFID 886 899 BY SIMILARITY.
 FT DISULFID 905 917 BY SIMILARITY.
 FT DISULFID 913 926 BY SIMILARITY.
 FT DISULFID 928 941 BY SIMILARITY.
 FT DISULFID 947 956 BY SIMILARITY.
 FT DISULFID 952 965 BY SIMILARITY.
 FT DISULFID 967 980 BY SIMILARITY.
 FT DISULFID 986 998 BY SIMILARITY.
 FT DISULFID 994 1007 BY SIMILARITY.
 FT DISULFID 1009 1023 BY SIMILARITY.
 FT DISULFID 1029 1042 BY SIMILARITY.
 FT DISULFID 1036 1051 BY SIMILARITY.
 FT DISULFID 1056 1068 BY SIMILARITY.
 FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 507 507 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1035 1035 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1184 AA; 126543 MW; CA48490A55F9EC5D CRC64;

Query Match 11.8%; Score 362.5; DB 1; Length 1184;
 Best Local Similarity 30.7%; Pred. No. 5,6e-18;
 Matches 107; Conservative 35; Mismatches 149; Indels 57; Gaps 17;

OY 21 GNAASARHHGLLASAROPGVCHYGT-----KLACCYWMRRNSKG--VCEATCEPG 68
 DB 806 GNCVDINECTSLSEPCRFSCINTVSGYTCQRPMLICARGYHASDDAKCVNDECETG 865
 OY 69 ---GKFGCV---GPNKCRGFGYT---GKTCSDVNEGCMKP-RRCQRCVNTHSY 116
 DB 866 VHRGEGGVCHNLPGSYRCDCAGFORDAFGRGCI-DVNECWASPGRLCQHTCENTLGSY 924
 OY 117 KCFLSGHMLPDA-TCVNSRTCAMINQYSCEDTEEGPCLCPSSGRLAPNGRCDI 175
 DB 925 RCGSASGFLAADKRCEDVNECEAQRCSQCANITYGYCYC-RQGQLADEGHTCDI 983
 OY 176 DECAAGK-VICPVNRCVNFSGSYCKCHIGFELYISGRYDCIDINECTMSHTCSHHA 234
 DB 984 DECAQAGILCTF--RCILNPGSYOCACPEQGYMTFANGR-SCKVDICALGTNCSBAE 1040
 OY 235 NCFWQGSFK---KCKQGY-KGNGLCRSALPENSVEVLAPGTINDRIKILAHKNSM 290
 DB 1041 TCHNIGSFRLRPECPPNVQVSKTCERTCHDELECONSAPRI--THYQLNFGTGL 1097
 OY 291 KKKAKIKNVTPERT-PTPKVNLQPNYERIEIVSRGNSHGKKGNEE 337
 DB 1098 LVPAHIFRIGPAPAFGTGTALNT-----IKGNEE 1127

RESULT 10
 MTN2_MOUSE
 ID MTN2_MOUSE STANDARD; PRT; 956 AA.
 AC 008746;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Matrilin-2 precursor.
 GN MATN2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Eulastomus;
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BA1B/C; TISSUE=Limb;
 RX MEDLINE=97238863; PubMed=9083061;
 RA Deak F., Plescha D., Bachrati C., Paulsson M., Kiss I.;
 RT Primary structure and expression of matrilin-2, the closest relative
 RT of cartilage matrix protein within the von Willebrand factor type A-
 RT like module superfamily.";
 RL J. Biol. Chem. 272:9268-9274(1997).

CC -1- FUNCTION: INVOLVED IN MATRIX ASSEMBLY (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: DETECTED IN A VARIETY OF ORGANS, INCLUDING
 CC CALVARIA, UTERUS, HEART, AND BRAIN, AS WELL AS FIBROBLAST AND
 CC OSTEOBLAST CELL LINES.
 CC -1- SIMILARITY: CONTAINS 10 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 VWFA DOMAINS.
 CC -----
 CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U69262; AAC53163.1; -;
 DR HSSP: P35555; IEMN.
 DR MGD: MGI:109613; Matn2.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF-Ca.
 DR InterPro: IPR002035; VWFA.
 DR Pfam: PF00008; EGF; 10.
 DR Pfam: PF00092; Vwa; 2.
 DR PRINTS: PR00453; VWFADOMAIN.
 DR SMART: SM00181; EGF; 10.
 DR SMART: SM00327; VWA; 2.
 DR PROSITE: PS00010; ASX_HYDROXYL; 7.
 DR PROSITE: PS01186; EGF; 2; 9.
 DR PROSITE: PS0234; VWFA; 2.
 DR EGF-like domain; Signal; Glycoprotein; Repeat; Coiled coil.
 KW EGF-like domain; Signal; Glycoprotein; Repeat; Coiled coil.
 FT CHAIN 1 23
 FT SIGNAL 1 23
 FT CHAIN 24 956
 FT DOMAIN 57 232
 FT DOMAIN 238 278
 FT DOMAIN 279 319
 FT DOMAIN 320 360
 FT DOMAIN 361 401
 FT DOMAIN 402 442
 FT DOMAIN 443 483
 FT DOMAIN 484 524
 FT DOMAIN 525 565
 FT DOMAIN 566 606
 FT DOMAIN 607 647
 FT DOMAIN 647 687
 FT DOMAIN 687 727
 FT DOMAIN 727 767
 FT DOMAIN 767 807
 FT DOMAIN 807 847
 FT DOMAIN 847 887
 FT DOMAIN 887 927
 FT DOMAIN 927 967
 FT DOMAIN 967 1007
 FT DOMAIN 1007 1047
 FT DOMAIN 1047 1087
 FT DOMAIN 1087 1127
 FT DOMAIN 1127 1167
 FT DOMAIN 1167 1207
 FT DOMAIN 1207 1247
 FT DOMAIN 1247 1287
 FT DOMAIN 1287 1327
 FT DOMAIN 1327 1367
 FT DOMAIN 1367 1407
 FT DOMAIN 1407 1447
 FT DOMAIN 1447 1487
 FT DOMAIN 1487 1527
 FT DOMAIN 1527 1567
 FT DOMAIN 1567 1607
 FT DOMAIN 1607 1647
 FT DOMAIN 1647 1687
 FT DOMAIN 1687 1727
 FT DOMAIN 1727 1767
 FT DOMAIN 1767 1807
 FT DOMAIN 1807 1847
 FT DOMAIN 1847 1887
 FT DOMAIN 1887 1927
 FT DOMAIN 1927 1967
 FT DOMAIN 1967 2007
 FT DOMAIN 2007 2047
 FT DOMAIN 2047 2087
 FT DOMAIN 2087 2127
 FT DOMAIN 2127 2167
 FT DOMAIN 2167 2207
 FT DOMAIN 2207 2247
 FT DOMAIN 2247 2287
 FT DOMAIN 2287 2327
 FT DOMAIN 2327 2367
 FT DOMAIN 2367 2407
 FT DOMAIN 2407 2447
 FT DOMAIN 2447 2487
 FT DOMAIN 2487 2527
 FT DOMAIN 2527 2567
 FT DOMAIN 2567 2607
 FT DOMAIN 2607 2647
 FT DOMAIN 2647 2687
 FT DOMAIN 2687 2727
 FT DOMAIN 2727 2767
 FT DOMAIN 2767 2807
 FT DOMAIN 2807 2847
 FT DOMAIN 2847 2887
 FT DOMAIN 2887 2927
 FT DOMAIN 2927 2967
 FT DOMAIN 2967 3007
 FT DOMAIN 3007 3047
 FT DOMAIN 3047 3087
 FT DOMAIN 3087 3127
 FT DOMAIN 3127 3167
 FT DOMAIN 3167 3207
 FT DOMAIN 3207 3247
 FT DOMAIN 3247 3287
 FT DOMAIN 3287 3327
 FT DOMAIN 3327 3367
 FT DOMAIN 3367 3407
 FT DOMAIN 3407 3447
 FT DOMAIN 3447 3487
 FT DOMAIN 3487 3527
 FT DOMAIN 3527 3567
 FT DOMAIN 3567 3607
 FT DOMAIN 3607 3647
 FT DOMAIN 3647 3687
 FT DOMAIN 3687 3727
 FT DOMAIN 3727 3767
 FT DOMAIN 3767 3807
 FT DOMAIN 3807 3847
 FT DOMAIN 3847 3887
 FT DOMAIN 3887 3927
 FT DOMAIN 3927 3967
 FT DOMAIN 3967 4007
 FT DOMAIN 4007 4047
 FT DOMAIN 4047 4087
 FT DOMAIN 4087 4127
 FT DOMAIN 4127 4167
 FT DOMAIN 4167 4207
 FT DOMAIN 4207 4247
 FT DOMAIN 4247 4287
 FT DOMAIN 4287 4327
 FT DOMAIN 4327 4367
 FT DOMAIN 4367 4407
 FT DOMAIN 4407 4447
 FT DOMAIN 4447 4487
 FT DOMAIN 4487 4527
 FT DOMAIN 4527 4567
 FT DOMAIN 4567 4607
 FT DOMAIN 4607 4647
 FT DOMAIN 4647 4687
 FT DOMAIN 4687 4727
 FT DOMAIN 4727 4767
 FT DOMAIN 4767 4807
 FT DOMAIN 4807 4847
 FT DOMAIN 4847 4887
 FT DOMAIN 4887 4927
 FT DOMAIN 4927 4967
 FT DOMAIN 4967 5007
 FT DOMAIN 5007 5047
 FT DOMAIN 5047 5087
 FT DOMAIN 5087 5127
 FT DOMAIN 5127 5167
 FT DOMAIN 5167 5207
 FT DOMAIN 5207 5247
 FT DOMAIN 5247 5287
 FT DOMAIN 5287 5327
 FT DOMAIN 5327 5367
 FT DOMAIN 5367 5407
 FT DOMAIN 5407 5447
 FT DOMAIN 5447 5487
 FT DOMAIN 5487 5527
 FT DOMAIN 5527 5567
 FT DOMAIN 5567 5607
 FT DOMAIN 5607 5647
 FT DOMAIN 5647 5687
 FT DOMAIN 5687 5727
 FT DOMAIN 5727 5767
 FT DOMAIN 5767 5807
 FT DOMAIN 5807 5847
 FT DOMAIN 5847 5887
 FT DOMAIN 5887 5927
 FT DOMAIN 5927 5967
 FT DOMAIN 5967 6007
 FT DOMAIN 6007 6047
 FT DOMAIN 6047 6087
 FT DOMAIN 6087 6127
 FT DOMAIN 6127 6167
 FT DOMAIN 6167 6207
 FT DOMAIN 6207 6247
 FT DOMAIN 6247 6287
 FT DOMAIN 6287 6327
 FT DOMAIN 6327 6367
 FT DOMAIN 6367 6407
 FT DOMAIN 6407 6447
 FT DOMAIN 6447 6487
 FT DOMAIN 6487 6527
 FT DOMAIN 6527 6567
 FT DOMAIN 6567 6607
 FT DOMAIN 6607 6647
 FT DOMAIN 6647 6687
 FT DOMAIN 6687 6727
 FT DOMAIN 6727 6767
 FT DOMAIN 6767 6807
 FT DOMAIN 6807 6847
 FT DOMAIN 6847 6887
 FT DOMAIN 6887 6927
 FT DOMAIN 6927 6967
 FT DOMAIN 6967 7007
 FT DOMAIN 7007 7047
 FT DOMAIN 7047 7087
 FT DOMAIN 7087 7127
 FT DOMAIN 7127 7167
 FT DOMAIN 7167 7207
 FT DOMAIN 7207 7247
 FT DOMAIN 7247 7287
 FT DOMAIN 7287 7327
 FT DOMAIN 7327 7367
 FT DOMAIN 7367 7407
 FT DOMAIN 7407 7447
 FT DOMAIN 7447 7487
 FT DOMAIN 7487 7527
 FT DOMAIN 7527 7567
 FT DOMAIN 7567 7607
 FT DOMAIN 7607 7647
 FT DOMAIN 7647 7687
 FT DOMAIN 7687 7727
 FT DOMAIN 7727 7767
 FT DOMAIN 7767 7807
 FT DOMAIN 7807 7847
 FT DOMAIN 7847 7887
 FT DOMAIN 7887 7927
 FT DOMAIN 7927 7967
 FT DOMAIN 7967 8007
 FT DOMAIN 8007 8047
 FT DOMAIN 8047 8087
 FT DOMAIN 8087 8127
 FT DOMAIN 8127 8167
 FT DOMAIN 8167 8207
 FT DOMAIN 8207 8247
 FT DOMAIN 8247 8287
 FT DOMAIN 8287 8327
 FT DOMAIN 8327 8367
 FT DOMAIN 8367 8407
 FT DOMAIN 8407 8447
 FT DOMAIN 8447 8487
 FT DOMAIN 8487 8527
 FT DOMAIN 8527 8567
 FT DOMAIN 8567 8607
 FT DOMAIN 8607 8647
 FT DOMAIN 8647 8687
 FT DOMAIN 8687 8727
 FT DOMAIN 8727 8767
 FT DOMAIN 8767 8807
 FT DOMAIN 8807 8847
 FT DOMAIN 8847 8887
 FT DOMAIN 8887 8927
 FT DOMAIN 8927 8967
 FT DOMAIN 8967 9007
 FT DOMAIN 9007 9047
 FT DOMAIN 9047 9087
 FT DOMAIN 9087 9127
 FT DOMAIN 9127 9167
 FT DOMAIN 9167 9207
 FT DOMAIN 9207 9247
 FT DOMAIN 9247 9287
 FT DOMAIN 9287 9327
 FT DOMAIN 9327 9367
 FT DOMAIN 9367 9407
 FT DOMAIN 9407 9447
 FT DOMAIN 9447 9487
 FT DOMAIN 9487 9527
 FT DOMAIN 9527 9567
 FT DOMAIN 9567 9607
 FT DOMAIN 9607 9647
 FT DOMAIN 9647 9687
 FT DOMAIN 9687 9727
 FT DOMAIN 9727 9767
 FT DOMAIN 9767 9807
 FT DOMAIN 9807 9847
 FT DOMAIN 9847 9887
 FT DOMAIN 9887 9927
 FT DOMAIN 9927 9967
 FT DOMAIN 9967 10007
 FT DOMAIN 10007 10047
 FT DOMAIN 10047 10087
 FT DOMAIN 10087 10127
 FT DOMAIN 10127 10167
 FT DOMAIN 10167 10207
 FT DOMAIN 10207 10247
 FT DOMAIN 10247 10287
 FT DOMAIN 10287 10327
 FT DOMAIN 10327 10367
 FT DOMAIN 10367 10407
 FT DOMAIN 10407 10447
 FT DOMAIN 10447 10487
 FT DOMAIN 10487 10527
 FT DOMAIN 10527 10567
 FT DOMAIN 10567 10607
 FT DOMAIN 10607 10647
 FT DOMAIN 10647 10687
 FT DOMAIN 10687 10727
 FT DOMAIN 10727 10767
 FT DOMAIN 10767 10807
 FT DOMAIN 10807 10847
 FT DOMAIN 10847 10887
 FT DOMAIN 10887 10927
 FT DOMAIN 10927 10967
 FT DOMAIN 10967 11007
 FT DOMAIN 11007 11047
 FT DOMAIN 11047 11087
 FT DOMAIN 11087 11127
 FT DOMAIN 11127 11167
 FT DOMAIN 11167 11207
 FT DOMAIN 11207 11247
 FT DOMAIN 11247 11287
 FT DOMAIN 11287 11327
 FT DOMAIN 11327 11367
 FT DOMAIN 11367 11407
 FT DOMAIN 11407 11447
 FT DOMAIN 11447 11487
 FT DOMAIN 11487 11527
 FT DOMAIN 11527 11567
 FT DOMAIN 11567 11607
 FT DOMAIN 11607 11647
 FT DOMAIN 11647 11687
 FT DOMAIN 11687 11727
 FT DOMAIN 11727 11767
 FT DOMAIN 11767 11807
 FT DOMAIN 11807 11847
 FT DOMAIN 11847 11887
 FT DOMAIN 11887 11927
 FT DOMAIN 11927 11967
 FT DOMAIN 11967 12007
 FT DOMAIN 12007 12047
 FT DOMAIN 12047 12087
 FT DOMAIN 12087 12127
 FT DOMAIN 12127 12167
 FT DOMAIN 12167 12207
 FT DOMAIN 12207 12247
 FT DOMAIN 12247 12287
 FT DOMAIN 12287 12327
 FT DOMAIN 12327 12367
 FT DOMAIN 12367 12407
 FT DOMAIN 12407 12447
 FT DOMAIN 12447 12487
 FT DOMAIN 12487 12527
 FT DOMAIN 12527 12567
 FT DOMAIN 12567 12607
 FT DOMAIN 12607 12647
 FT DOMAIN 12647 12687
 FT DOMAIN 12687 12727
 FT DOMAIN 12727 12767
 FT DOMAIN 12767 12807
 FT DOMAIN 12807 12847
 FT DOMAIN 12847 12887
 FT DOMAIN 12887 12927
 FT DOMAIN 12927 12967
 FT DOMAIN 12967 13007
 FT DOMAIN 13007 13047
 FT DOMAIN 13047 13087
 FT DOMAIN 13087 13127
 FT DOMAIN 13127 13167
 FT DOMAIN 13167 13207
 FT DOMAIN 13207 13247
 FT DOMAIN 13247 13287
 FT DOMAIN 13287 13327
 FT DOMAIN 13327 13367
 FT DOMAIN 13367 13407
 FT DOMAIN 13407 13447
 FT DOMAIN 13447 13487
 FT DOMAIN 13487 13527
 FT DOMAIN 13527 13567
 FT DOMAIN 13567 13607
 FT DOMAIN 13607 13647
 FT DOMAIN 13647 13687
 FT DOMAIN 13687 13727
 FT DOMAIN 13727 13767
 FT DOMAIN 13767 13807
 FT DOMAIN 13807 13847
 FT DOMAIN 13847 13887
 FT DOMAIN 13887 13927
 FT DOMAIN 13927 13967
 FT DOMAIN 13967 14007
 FT DOMAIN 14007 14047
 FT DOMAIN 14047 14087
 FT DOMAIN 14087 14127
 FT DOMAIN 14127 14167
 FT DOMAIN 14167 14207
 FT DOMAIN 14207 14247
 FT DOMAIN 14247 14287
 FT DOMAIN 14287 14327
 FT DOMAIN 14327 14367
 FT DOMAIN 14367 14407
 FT DOMAIN 14407 14447
 FT DOMAIN 14447 14487
 FT DOMAIN 14487 14527
 FT DOMAIN 14527 14567
 FT DOMAIN 14567 14607
 FT DOMAIN 14607 14647
 FT DOMAIN 14647 14687
 FT DOMAIN 14687 14727
 FT DOMAIN 14727 14767
 FT DOMAIN 14767 14807
 FT DOMAIN 14807 14847
 FT DOMAIN 14847 14887
 FT DOMAIN 14887 14927
 FT DOMAIN 14927 14967
 FT DOMAIN 14967 15007
 FT DOMAIN 15007 15047
 FT DOMAIN 15047 15087
 FT DOMAIN 15087 15127
 FT DOMAIN 15127 15167
 FT DOMAIN 15167 15207
 FT DOMAIN 15207 15247
 FT DOMAIN 15247 15287
 FT DOMAIN 15287 15327
 FT DOMAIN 15327 15367
 FT DOMAIN 15367 15407
 FT DOMAIN 15407 15447
 FT DOMAIN 15447 15487
 FT DOMAIN 15487 15527
 FT DOMAIN 15527 15567
 FT DOMAIN 15567 15607
 FT DOMAIN 15607 15647
 FT DOMAIN 15647 15687
 FT DOMAIN 15687 15727
 FT DOMAIN 15727 15767
 FT DOMAIN 15767 15807
 FT DOMAIN 15807 15847
 FT DOMAIN 15847 15887
 FT DOMAIN 15887 15927
 FT DOMAIN 15927 15967
 FT DOMAIN 15967 16007
 FT DOMAIN 16007 16047
 FT DOMAIN 16047 16087
 FT DOMAIN 16087 16127
 FT DOMAIN 16127 16167
 FT DOMAIN 16167 16207
 FT DOMAIN 16207 16247
 FT DOMAIN 16247 16287
 FT DOMAIN 16287 16327
 FT DOMAIN 16327 16367
 FT DOMAIN 16367 16407
 FT DOMAIN 16407 16447
 FT DOMAIN 16447 16487
 FT DOMAIN 16487 16527
 FT DOMAIN 16527 16567
 FT DOMAIN 16567 16607
 FT DOMAIN 16607 16647
 FT DOMAIN 16647 16687
 FT DOMAIN 16687 16727
 FT DOMAIN 16727 16767
 FT DOMAIN 16767 16807
 FT DOMAIN 16807 16847
 FT DOMAIN 16847 16887
 FT DOMAIN 16887 16927
 FT DOMAIN 16927 16967
 FT DOMAIN 16967 17007
 FT DOMAIN 17007 17047
 FT DOMAIN 17047 17087
 FT DOMAIN 17087 17127
 FT DOMAIN 17127 17167
 FT DOMAIN 17167 17207
 FT DOMAIN 17207 17247
 FT DOMAIN 17247 17287
 FT DOMAIN 17287 17327
 FT DOMAIN 17327 17367
 FT DOMAIN 17367 17407
 FT DOMAIN 17407 17447
 FT DOMAIN 17447 17487
 FT DOMAIN 17487 17527
 FT DOMAIN 17527 17567
 FT DOMAIN 17567 17607
 FT DOMAIN 17607 17647
 FT DOMAIN 17647 17687
 FT DOMAIN 17687 17727
 FT DOMAIN 17727 17767
 FT DOMAIN 17767 17807
 FT DOMAIN 17807 17847
 FT DOMAIN 17847 17887
 FT DOMAIN 17887 17927
 FT DOMAIN 17927 17967
 FT DOMAIN 17967 18007
 FT DOMAIN 18007 18047
 FT DOMAIN 18047 18087
 FT DOMAIN 18087 18127
 FT DOMAIN 18127 18167
 FT DOMAIN 18167 18207
 FT DOMAIN 18207 18247
 FT DOMAIN 18247 18287
 FT DOMAIN 18287 18327
 FT DOMAIN 18327 18367
 FT DOMAIN 18367 18407
 FT DOMAIN 18407 18447
 FT DOMAIN 18447 18487
 FT DOMAIN 18487 18527
 FT DOMAIN 18527 18567
 FT DOMAIN 18567 18607
 FT DOMAIN 18607 18647
 FT DOMAIN 18647 18687
 FT DOMAIN 18687 18727
 FT DOMAIN 18727 18767
 FT DOMAIN 18767 18807
 FT DOMAIN 18807 18847
 FT DOMAIN 18847 18887
 FT DOMAIN 18887 18927
 FT DOMAIN 18927 18967
 FT DOMAIN 18967 19007
 FT DOMAIN 19007 19047
 FT DOMAIN 19047 19087
 FT DOMAIN 19087 19127
 FT DOMAIN 19127 19167
 FT DOMAIN 19167 19207
 FT DOMAIN 19207 19247
 FT DOMAIN 19247 19287
 FT DOMAIN 19287 19327
 FT DOMAIN 19327 19367
 FT DOMAIN 19367 19407
 FT DOMAIN 19407 19447
 FT DOMAIN 19447 19487
 FT DOMAIN 19487 19527
 FT DOMAIN 19527 19567
 FT DOMAIN 19567 19607
 FT DOMAIN 19607 19647
 FT DOMAIN 19647 19687
 FT DOMAIN 19687 19727
 FT DOMAIN 19727 19767
 FT DOMAIN 19767 19807
 FT DOMAIN 19807 19847
 FT DOMAIN 19847 19887
 FT DOMAIN 19887 19927
 FT DOMAIN 19927 19967
 FT DOMAIN 19967 20007
 FT DOMAIN 20007 20047
 FT DOMAIN 20047 20087
 FT DOMAIN 20087 20127
 FT DOMAIN 20127 20167
 FT DOMAIN 20167 20207
 FT DOMAIN 20207 20247
 FT DOMAIN 20247 20287
 FT DOMAIN 20287 20327
 FT DOMAIN 20327 20367
 FT DOMAIN 20367 20407
 FT DOMAIN 20407 20447
 FT DOMAIN 20447 20487
 FT DOMAIN 20487 20527
 FT DOMAIN 20527 20567
 FT DOMAIN 20567 20607
 FT DOMAIN 20607 20647
 FT DOMAIN 20647 20687
 FT DOMAIN 20687 20727
 FT DOMAIN 20727 20767
 FT DOMAIN 20767 20807
 FT DOMAIN 20807 20847
 FT DOMAIN 20847 20887
 FT DOMAIN 20887 20927
 FT DOMAIN 20927 20967
 FT DOMAIN 20967 21007
 FT DOMAIN 21007 21047
 FT DOMAIN 21047 21087
 FT DOMAIN 21087 21127
 FT DOMAIN 21127 21167
 FT DOMAIN 21167 21207
 FT DOMAIN 21207 21247
 FT DOMAIN 21247 21287
 FT DOMAIN 21287 21327
 FT DOMAIN 21327 21367
 FT DOMAIN 21367 21407
 FT DOMAIN 21407 21447
 FT DOMAIN 21447 21487
 FT DOMAIN 21487 21527
 FT DOMAIN 21527 21567
 FT DOMAIN 21567 21607
 FT DOMAIN 21607 21647
 FT DOMAIN 21647 21687
 FT DOMAIN 21687 21727
 FT DOMAIN 21727 21767
 FT DOMAIN 21767 21807
 FT DOMAIN 21807 21847
 FT DOMAIN 21847 21887
 FT DOMAIN 21887 21927
 FT DOMAIN 21927 21967
 FT DOMAIN 21967 22007
 FT DOMAIN 22007 22047
 FT DOMAIN 22047 22087
 FT DOMAIN 22087 22127
 FT DOMAIN 22127 22167
 FT DOMAIN 22167 22207
 FT DOMAIN 22207 22247
 FT DOMAIN 22247 22287
 FT DOMAIN 22287 22327
 FT DOMAIN 22327 22367
 FT DOMAIN 22367 22407
 FT DOMAIN 22407 22447
 FT DOMAIN 22447 22487
 FT DOMAIN 22487 22527
 FT DOMAIN 22527 22567
 FT DOMAIN 22567 22607
 FT DOMAIN 22607 22647
 FT DOMAIN 22647 22687
 FT DOMAIN 22687 22727
 FT DOMAIN 22727 22767
 FT DOMAIN 22767 22807
 FT DOMAIN 22807 22847
 FT DOMAIN 22847 22887
 FT DOMAIN 22887 22927
 FT DOMAIN 22927 22967
 FT DOMAIN 22967 23007
 FT DOMAIN 23007 23047
 FT DOMAIN 23047 23087
 FT DOMAIN 23087 2

```

FT DISULFID 592 605 BY SIMILARITY.
FT DISULFID 611 622 BY SIMILARITY.
FT DISULFID 618 631 BY SIMILARITY.
FT DISULFID 633 646 BY SIMILARITY.
FT CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 890 890 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 956 AA; 106779 MW; 3EAC22770B6E6BEE CRC64;

Query Match 11.18; Score 339; DB 1; Length 956;
Best Local Similarity 23.28; Pred. No. 2e-16;
Matches 140; Conservative 77; Mismatches 225; Indels 162; Gaps 28;

OY 25 SARHGL---LASAROPCVCHYGTKLACCYGMRN-SKGYE-----ATCEPGCKFGCV 75
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 366 ASSNHGCOHECVNMQTSALCR-----CLKFMLNPDKTKRIRNYCLNKGCEH-ECV 418
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 76 GPNK---CRCPGY---TGFTCSODVNECGMKRPPQCHRCVNTHGSKYKFCGLGMLMP 128
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 419 NTEGHCRCRGOYNLDENGKTSR-VDHCAQODHCEQLCLNTEESFVCCSEGLIND 477
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 129 D-ATCVNSRTCAMIN--QYSCDTEBGPQLCPSSGLRLAPNGDCLDIDECASGKYIC 185
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 478 DLKTCRADYCLLSNHGCEYCVMTKSFACQCP-EGHVLNSDGTAKLIDSCALGDHGC 536
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 186 PYN-----RR-----CVNFGSYCKCHIGF 206
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 537 EHSVSSFSDFVCOCFEYILRDGKTCRRKDYCDVNHGCEHLGVNCGESTYCKCLEGF 596
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 207 ELQYISGRYDCIDINECTMDSHSCSHNANCENTOGSFKCKQGY--KGNGLRCSAIPEN 264
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 597 RLADGKR--CRKNVCKSTQHGCEH--MCVNGNSYLCRCSEGVLAEDGKHCKRCEG 652
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 265 SYKEVLAPGCIKIRIKKLLAHKNSMKKAKIKVTPETRP--TPKYNLDQPNTE---E 320
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 653 PIDLFWIDGS-----KSLSEENFETVKHFVTGIIDSLAVSPKARGLQYQVFRTE 706
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 321 IVSRGNSHGSKGKNEKMEKLEDEKREK-----ALKNDIEBRSLRGVFFPKVNEA 374
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 707 FTLNGFS-----AKMKKAVTHMKYMGKSGMTGLAKIMFESLTFQVEGARPPSTQV 759
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 375 GEFGLLVQ-----RKALTSKLEH-----KDNISVD 401
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 760 PRVAIVFDGRADVDSEWASKAKANGITMYAVGVKAIEELOEIASSEPIDKHIFYAED 819
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 402 CS-----FNHGICDMKOD---REDDFDNPNPADRONAIGFYAVPRLAGKKKIGILK 450
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 820 FSTWGEISEKLEGEICALEDSGGKODSAWD-----LPQAHQPTPEPEPV 866
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 451 LLLPDLQPSNFCU---LFDYRLAGDVKGLRVFKNSNNALAMEKTTSEDEKKTGKI 506
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 867 IKINDLSCSNFAVQHRFLFEEDNLSRSTOKLFHSTKSSGNPL-----EESQDQCKENTL 921
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 507 QLYQ 510
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 922 ILFQ 925

RESULT 11
FBL_HUMAN STANDARD; PRT; 703 AA.
AC P23142; P23143; P23144; P37888; Q9UGR4;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fbulin-1 precursor.
GN FBLN1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
RX MEDLINE=91100426; PubMed=2269669;

```

```

RA Argaves W.S., Tran H., Burgess W.H., Dickerson K.;
RT "Fibulin is an extracellular matrix and plasma glycoprotein with
RL repeated domain structure."
RL J. Cell Biol. 111:3155-3164(1990).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM D).
RA Argaves S.;
RL Submitted (SEP-1993) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE OF 567-703 FROM N.A. (ISOFORM C).
RA Connor R.;
RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE OF 30-44.
RX MEDLINE=89354537; PubMed=2527614;
RA Argaves W.S., Dickerson K., Burgess W.H., Ruoslahti E.;
RT "Fibulin, a novel protein that interacts with the fibronectin
RL receptor beta subunit cytoplasmic domain."
RL Cell 58:623-629(1989).
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX.
CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS: A, B, C AND D (SHOWN HERE); ARE
CC PRODUCED BY ALTERNATIVE SPLICING. THEY DIFFER ONLY IN THEIR C-
CC TERMINAL REGIONS.
CC -1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib.ch).
CC
DR EMBL; X53741; CA37770.1; -
DR EMBL; X53742; CA37771.1; -
DR EMBL; X53743; CA37772.1; -
DR EMBL; U01244; AAB17099.1; -
DR EMBL; 295331; CAB62960.1; -
DR PIR; A32826; A32826.
DR PIR; A36346; A36346.
DR PIR; B36346; B36346.
DR PIR; C36346; C36346.
DR HSSP; P35555; 1EMN.
DR MIM; 135820; -
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR Pfam; PF01821; ANATO; 3.
DR Pfam; PF00008; EGF; 6.
DR SMART; SM00104; ANATO; 3.
DR SMART; SM00179; EGF_CA; 7.
DR SMART; SM00001; EGF_Like; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS00022; EGF_1; FALSE NEG.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; 3.
DR PROSITE; PS01178; ANAPHYLATOXIN_2; 3.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA; 8.
DR Signal; Alternative splicing: Glycoprotein; Extracellular matrix;
KW Repeat; EGF-like domain; Calcium-binding.
FT SIGNAL 1 29
FT CHAIN 30 703 FIBULIN-1.
FT DOMAIN 36 76 ANAPHYLATOXIN-LIKE 1.
FT DOMAIN 77 111 ANAPHYLATOXIN-LIKE 2.
FT DOMAIN 112 144 ANAPHYLATOXIN-LIKE 3.
FT DOMAIN 176 215 EGF-LIKE 1.
FT DOMAIN 216 261 EGF-LIKE 2.
FT DOMAIN 262 307 CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 308 355 EGF-LIKE 3.
FT DOMAIN 356 398 EGF-LIKE 4.
FT DOMAIN 399 440 EGF-LIKE 5.
FT DOMAIN EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).

```



```

FT DOMAIN 441 480 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 481 524 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 525 578 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT DISULFID 36 61 BY SIMILARITY.
FT DISULFID 37 68 BY SIMILARITY.
FT DISULFID 50 69 BY SIMILARITY.
FT DISULFID 78 109 BY SIMILARITY.
FT DISULFID 91 110 BY SIMILARITY.
FT DISULFID 112 136 BY SIMILARITY.
FT DISULFID 113 143 BY SIMILARITY.
FT DISULFID 126 144 BY SIMILARITY.
FT DISULFID 180 190 BY SIMILARITY.
FT DISULFID 186 199 BY SIMILARITY.
FT DISULFID 201 214 BY SIMILARITY.
FT DISULFID 220 233 BY SIMILARITY.
FT DISULFID 227 242 BY SIMILARITY.
FT DISULFID 248 260 BY SIMILARITY.
FT DISULFID 266 279 BY SIMILARITY.
FT DISULFID 273 288 BY SIMILARITY.
FT DISULFID 294 306 BY SIMILARITY.
FT DISULFID 312 325 BY SIMILARITY.
FT DISULFID 319 334 BY SIMILARITY.
FT DISULFID 341 354 BY SIMILARITY.
FT DISULFID 360 373 BY SIMILARITY.
FT DISULFID 367 382 BY SIMILARITY.
FT DISULFID 384 397 BY SIMILARITY.
FT DISULFID 403 415 BY SIMILARITY.
FT DISULFID 411 424 BY SIMILARITY.
FT DISULFID 426 439 BY SIMILARITY.
FT DISULFID 445 454 BY SIMILARITY.
FT DISULFID 450 463 BY SIMILARITY.
FT DISULFID 465 479 BY SIMILARITY.
FT DISULFID 485 498 BY SIMILARITY.
FT DISULFID 494 507 BY SIMILARITY.
FT DISULFID 509 523 BY SIMILARITY.
FT DISULFID 529 542 BY SIMILARITY.
FT DISULFID 536 551 BY SIMILARITY.
FT DISULFID 556 577 BY SIMILARITY.
FT CARBOHYD 98 577 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 535 535 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 539 539 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VASAPLIC 567 703 MISSING (IN ISOPFORM A).
FT VASAPLIC 567 601 LQKERTDPRCKTSCRPNDVTCVDPVHTTSHYI -> QK
FT VASAPLIC 567 601 SKGRONTNPAAGSSKEDCVLWKKGLEPHTLDA (IN
FT VASAPLIC 567 703 ISOPFORM B).
FT VASAPLIC 567 703 LQKERTDPRCKTSCRPNDVTCVDPVHTTSHYI
FT VASAPLIC 567 703 EFTRPDEITFLRATPPHAPASOANIIFDITGNLSDPDI
FT VASAPLIC 567 703 KRYMDGMYGVROVRPIVGPENHATKEMNYVGVGVSHR
FT VASAPLIC 567 703 NVNVRVPESEYR -> RCRRLCHEMRECKSLRRTYX
FT VASAPLIC 567 703 HLRFPNIQAPAVVFRMGPSAVFEDSMQLITGNEGEFF
FT VASAPLIC 567 703 TTRKRVSPHSGVALTKPVEPERDLITVYKMDLSRGTVSSF
FT VASAPLIC 567 703 VAKLFIYSAEL (IN ISOPFORM C).
FT VASAPLIC 567 703 C -> S (IN REF. 4).
FT VASAPLIC 567 703 HR -> SH (IN REF. 4).
FT VASAPLIC 567 703 EFD8465BA2D3A25 CMC64.
SO SEQUENCE 703 AA: 77274 MW: 77274 MW: EFD8465BA2D3A25 CMC64.

Query Match 11.0%, Score 338; DB 1; Length 703;
Best Local Similarity 34.9%, Pred. No. 1.7e-16;
Matches 88; Conservative 30; Mismatches 86; Indels 48; Gaps 15;

```

```

Db 460 SYQCCRRRQYOLSDVDG-VTCEDIDECALPTGHCITCY--RCINIPGSPQSCSSPGYRL 516
OY 253 -GNGLRCSAIPR 263
Db 517 APNGRNCODIDE 528

RESULT 12
NOTC_DROME STANDARD: PRT: 2703 AA.
ID NOTC_DROME STANDARD: PRT: 2703 AA.
AC P07207; P04154; O97458; Q9W4T8.
AT 01-NOV-1986 (Rel. 03, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Neurogenic locus Notch protein precursor.
OS N OR EG:140G11.1 OR EG:163A10.2 OR CG33936.
OC Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Oregon-R; TISSUE-Embryo;
RX MEDLINE=86079539; PubMed=3935325;
RA Wharton K.A., Johansen K.M., Xu T., Artavanis-Tsakonas S.;
RT "Nucleotide sequence from the neurogenic locus notch implies a gene
RT product that shares homology with proteins containing EGF-like
RT repeats."
RL Cell 43:567-581(1985).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Canton-S; and Oregon-R; TISSUE-Embryo;
RX MEDLINE=87064624; PubMed=3097517;
RA Kidd S., Kelley M.R., Young M.W.;
RT "Sequence of the notch locus of Drosophila melanogaster: relationship
RT of the encoded protein to mammalian clotting and growth factors."
RL Mol. Cell. Biol. 6:3094-3108(1986).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle G., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abtill J.F., Agbayani A., An H.-J., Andrews-Plannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Beous P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cayley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dudin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Gloder A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.D., Wei M.-H., Ibegwam C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Iel Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., Meleod M.P., McPherson D.,
RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

```


RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissensbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveli J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Glibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*,"
 RL Science 287:2185-2195(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Oregon-R.
 RX MEDLINE=20196011; PubMed=10731137;
 RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,
 RA Barrell B.G., Ferraz C., Vidal S., Brun C., Demallies J., Cadieu E.,
 RA Dreano S., Gloux S., Leleau V., Mottier S., Galibert F., Borkova D.,
 RA Milana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,
 RA Papadogiannakis G., Spanos L., Cox S., Madueno E., de Pablos B.,
 RA Modolell J., Peter A., Schoettler P., Werner M., Moutikoti F.,
 RA Belinert N., Dove G., Schaefer U., Jaackie H., Bucheton A.,
 RA Callister D.M., Campbell L.A., Darlamsiou A., Henderson N.S.,
 RA McMillan P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C.,
 RA Glover D.M.;
 RT "From sequence to chromosome: the tip of the X chromosome of *D. melanogaster*,";
 RL Science 287:2220-2222(2000).
 RN [5]
 RP SEQUENCE OF 2505-2611 FROM N.A.
 RX MEDLINE=85099329; PubMed=2981631.
 RA Wharton K.A., Yedvobnick B., Flinerty V.G., Attavanis-Tsakonas S.;
 RT "opa: a novel family of transcribed repeats shared by the Notch locus
 RT and other developmentally regulated loci in *D. melanogaster*,";
 RL Cell 40:55-62(1985).
 RN [6]
 RP SEQUENCE OF 1-8 FROM N.A.
 RX MEDLINE=87257846; PubMed=3037327;
 RA Kelley M.R., Kidd S., Berg R.L., Young M.W.;
 RT "Restriction of P-element insertions at the Notch locus of *Drosophila melanogaster*,";
 RL Mol. Cell. Biol. 7:1545-1548(1987).
 RN [7]
 RP REVIEW.
 RA Harris W.A.;
 RT "Many cell types specified by Notch function,";
 RL Curr. Biol. 1:120-122(1991).
 CC -1- FUNCTION: NOTCH PROTEIN IS ESSENTIAL FOR PROPER DIFFERENTIATION OF ECTODERM.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- MISCELLANEOUS: SEPARATION OF NEUROBLASTS FROM THE ECTODERM INTO THE INNER PART OF EMBRYO IS ONE OF THE FIRST STEPS OF CNS DEVELOPMENT IN INSECTS, THIS PROCESS IS UNDER CONTROL OF THE NEUROGENIC GENES.
 CC -1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
 CC -1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
 CC -1- SIMILARITY: CONTAINS 6 ANK REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M16152; AAB59220.1; -;
 DR EMBL: M16153; AAB59220.1; JOINED.
 DR EMBL: M16149; AAB59220.1; JOINED.
 DR EMBL: M16150; AAB59220.1; JOINED.
 DR EMBL: M16151; AAB59220.1; JOINED.
 DR EMBL: K03508; AAA28725.1; -;
 DR EMBL: M13689; AAA28725.1; JOINED.
 DR EMBL: K03507; AAA28725.1; JOINED.
 DR EMBL: AE003426; AAF45848.2; -;

DR EMBL: AL035436; CAB37610.1; -;
 DR EMBL: AL035395; CAB37610.1; JOINED.
 DR EMBL: M12175; AAA74496.1; -;
 DR EMBL: M16025; AAA28726.1; -;
 DR PIR: A24420; A24420.
 DR PIR: A24768; A24768.
 DR PIR: A05267; A05267.
 DR HSSP: P00740; 11XA.
 DR FLYBase: FBgn0004647; N.
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000742; EGF_2.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR001438; EGF_II.
 DR InterPro: IPR000800; Notch.
 DR Pfam: PF00023; ank; 6.
 DR Pfam: PF00008; EGF; 36.
 DR Pfam: PF00066; notch; 3.
 DR PRINTS: PR00010; EGFBL00.
 DR PRINTS: PR01452; NOTCH.
 DR SMART: SM00248; ANK; 4.
 DR SMART: SM00179; EGF_CA; 23.
 DR SMART: SM00001; EGF_like; 13.
 DR SMART: SM00004; NL; 2.
 DR PROSITE: PS50088; ANK_REPEAT; 5.
 DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 22.
 DR PROSITE: PS00022; EGF_1; 34.
 DR PROSITE: PS01186; EGF_2; 28.
 DR PROSITE: PS01187; EGF_CA; 22.
 KW Differentiation: Neurogenesis; Repeat: ANK repeat; EGF-like domain; Transmembrane; Signal; Glycoprotein.
 KW SIGNAL 1 44
 FT CHAIN 45 2703
 FT DOMAIN 45 1745
 FT TRANSMM 1746 1766
 FT DOMAIN 1767 2703
 FT DOMAIN 58 95
 FT DOMAIN 96 136
 FT DOMAIN 139 176
 FT DOMAIN 177 215
 FT DOMAIN 217 253
 FT DOMAIN 255 291
 FT DOMAIN 293 329
 FT DOMAIN 331 370
 FT DOMAIN 372 408
 FT DOMAIN 409 447
 FT DOMAIN 449 486
 FT DOMAIN 488 524
 FT DOMAIN 524 562
 FT DOMAIN 564 600
 FT DOMAIN 602 637
 FT DOMAIN 639 675
 FT DOMAIN 677 713
 FT DOMAIN 715 751
 FT DOMAIN 753 789
 FT DOMAIN 791 827
 FT DOMAIN 829 865
 FT DOMAIN 867 905
 FT DOMAIN 907 944
 FT DOMAIN 946 982
 FT DOMAIN 984 1020
 FT DOMAIN 1022 1058
 Query Match 11.0%; Score 335.5; DB 1; Length 2703;
 Best Local Similarity 26.4%; Pred. No. 1,1e-15;
 Matches 96; Conservative 34; Mismatches 106; Indels 127; Gaps 17;
 OY 39 GYCHYGT-----KLACCYG--RRNSKGVCEFT-CEPGCKRECYGPKN--CRCPFGYTG 88
 DB 152 GYCHYGTLEEVYCAACANGYTCERCTKNLCASSPCRNAGATCTALAGSSSFTCCSPCPFTG 211

Matches 86; Conservative 31; Mismatches 87; Indels 48; Gaps 15;

QY 46 KLACVGRMRKSKVC-----EATCEGCKFGEVGNMKR-----CEPGV-----T 87
 Db 293 KLOCKSGFIQALNCIDINECLISAPCVGQFCINTGSGYTCQKNVPNGRHLNEE 352
 QY 88 KGTCSQDVNECGMKPRPC--OHRCVNTHSGYKFCFLSGHMMMPATCVNSRTGAMIN-- 142
 Db 353 GTRC-VVYDECAPPAEPCGKHGLNSGSRCECKAGFYF--DGI---SRTCVDINCO 406
 QY 143 -----CQYCEDETEBEGQCLCPSSGLRLANGRDCLDIDECASGKYICPNRRCVNTEG 196
 Db 407 RYPRRLCGHKCEKNTGSGFHCSC-SAGFRLSYDGRSCDEYNCLNS---PCSOECANVYG 461
 QY 197 SYVCKCHIGFELQYISGRYDIDINECTMDS--HTCSHANCFTMGSEFKCKG-KQGYK- 252
 Db 462 SYGCIYCRGIGLSDVDC-VTCEDIDECALPTGHICST--RCINIPSGFQSCPSGGRRL 518
 QY 253 -GNGLRCSAIP 263
 Db 519 APNGRNCQDIDE 530

RESULT 14
 NOTC_BRARE
 ID NOTC_BRARE STANDARD; PRT; 2437 AA.
 AC PA6530:
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Neurogenic locus notch homolog protein precursor.
 GN NOTCH OR NOTCH1A.
 OS Brachydanio rerio (zebrafish) (zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Cypriniformes; Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo; PubMed=8297791;
 RX MEDLINE=94128602; PubMed=8297791;
 RA Bierkamp C., Campos-Ortega J.A.;
 RT "A zebrafish homologue of the Drosophila neurogenic gene Notch and
 its pattern of transcription during early embryogenesis.";
 RL Mech. Dev. 43:87-100(1993).
 CC -1- FUNCTION: IMPLICATED IN CELL FATE SPECIFICATIONS DURING
 EMBRIO DEVELOPMENT. MAY BE INVOLVED IN THE FORMATION OF THE
 CC NEURAL PLATE, NOTOCHORD AND BRAIN VESICLES.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN ALL CELLS IN PREGASTRULATION
 CC STAGES. DURING GASTRULATION IS DIFFERENTIALLY EXPRESSED.
 CC ACCUMULATING PREDOMINANTLY IN THE PRECHORAL MESODERM AND
 CC NOTOCHORD. AT THE END OF GASTRULATION, EXPRESSED ALONG THE
 CC ANTERIOR-POSTERIOR AXIS INCLUDING THE DEVELOPING NEURAL PLATE
 CC AND DIFFERENTIATING MESODERM. ALSO PRESENT IN THE DEVELOPING
 CC BRAIN AND HEAD REGIONS.
 CC -1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.
 CC -1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
 CC -1- SIMILARITY: CONTAINS 6 ANK REPEATS.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X69088; CAA48831.1; -
 DR HSSP: P00740; JEDM.
 DR ZFIN: ZDB-GENE-990415-173; notchla.
 DR InterPro: IPR002110; ANK.

DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000742; EGF_2.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR001438; EGF_11.
 DR InterPro: IPR000800; Notch.
 DR Pfam: PF00023; ank; 6.
 DR Pfam: PF00008; EGF; 36.
 DR Pfam: PF00066; notch; 3.
 DR PRINTS: PR00010; EGFBLD.
 DR PRINTS: PR01452; NOTCH.
 DR SMART: SM00248; ANK; 5.
 DR SMART: SM00179; EGF_Ca; 19.
 DR SMART: SM00004; EGF-like; 16.
 DR SMART: SM00004; NL; 3.
 DR PROSITE: PS50088; ANK_REPEAT; 4.
 DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 23.
 DR PROSITE: PS00022; EGF_1; 34.
 DR PROSITE: PS01186; EGF_2; 28.
 DR PROSITE: PS01187; EGF_Ca; 22.
 KW Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;
 KW Transmembrane; Signal; Glycoprotein.
 KW SIGNAL
 FT CHAIN 21 2437
 FT DOMAIN 21 1724
 FT TRANSMEM 1725 1747
 FT DOMAIN 1748 2437
 FT DOMAIN 21 57
 FT DOMAIN 58 98
 FT DOMAIN 101 138
 FT DOMAIN 139 175
 FT DOMAIN 177 215
 FT DOMAIN 217 254
 FT DOMAIN 256 292
 FT DOMAIN 294 332
 FT DOMAIN 334 370
 FT DOMAIN 371 409
 FT DOMAIN 411 449
 FT DOMAIN 451 487
 FT DOMAIN 489 524
 FT DOMAIN 526 562
 FT DOMAIN 564 599
 FT DOMAIN 601 637
 FT DOMAIN 639 674
 FT DOMAIN 676 712
 FT DOMAIN 714 749
 FT DOMAIN 751 787
 FT DOMAIN 789 825
 FT DOMAIN 827 865
 FT DOMAIN 867 903
 FT DOMAIN 905 941
 FT DOMAIN 943 979
 FT DOMAIN 981 1017
 FT DOMAIN 1019 1055
 FT DOMAIN 1057 1093
 FT DOMAIN 1095 1141
 FT DOMAIN 1143 1179
 FT DOMAIN 1181 1217
 FT DOMAIN 1219 1263
 FT DOMAIN 1265 1303
 FT DOMAIN 1305 1344
 FT DOMAIN 1346 1382
 FT DOMAIN 1385 1423
 FT REPEAT 1446 1486
 FT REPEAT 1487 1520
 FT REPEAT 1521 1561
 FT REPEAT 1562 1591
 FT REPEAT 1592 1627
 FT REPEAT 1628 1667
 FT REPEAT 1668 1704
 FT REPEAT 1705 1741
 FT REPEAT 1742 1778
 FT REPEAT 1779 1815
 FT REPEAT 1816 1852
 FT REPEAT 1853 1889
 FT REPEAT 1890 1926
 FT REPEAT 1927 1963
 FT REPEAT 1964 2000
 FT REPEAT 2001 2037
 FT REPEAT 2038 2074
 FT REPEAT 2075 2111
 FT REPEAT 2112 2148
 FT REPEAT 2149 2185
 FT REPEAT 2186 2222
 FT REPEAT 2223 2259
 FT REPEAT 2260 2296
 FT REPEAT 2297 2333
 FT REPEAT 2334 2370
 FT REPEAT 2371 2407
 FT REPEAT 2408 2444
 FT REPEAT 2445 2481
 FT REPEAT 2482 2518
 FT REPEAT 2519 2555
 FT REPEAT 2556 2592
 FT REPEAT 2593 2629
 FT REPEAT 2630 2666
 FT REPEAT 2667 2703
 FT REPEAT 2704 2740
 FT REPEAT 2741 2777
 FT REPEAT 2778 2814
 FT REPEAT 2815 2851
 FT REPEAT 2852 2888
 FT REPEAT 2889 2925
 FT REPEAT 2926 2962
 FT REPEAT 2963 2999
 FT REPEAT 3000 3036
 FT REPEAT 3037 3073
 FT REPEAT 3074 3110
 FT REPEAT 3111 3147
 FT REPEAT 3148 3184
 FT REPEAT 3185 3221
 FT REPEAT 3222 3258
 FT REPEAT 3259 3295
 FT REPEAT 3296 3332
 FT REPEAT 3333 3369
 FT REPEAT 3370 3406
 FT REPEAT 3407 3443
 FT REPEAT 3444 3480
 FT REPEAT 3481 3517
 FT REPEAT 3518 3554
 FT REPEAT 3555 3591
 FT REPEAT 3592 3628
 FT REPEAT 3629 3665
 FT REPEAT 3666 3702
 FT REPEAT 3703 3739
 FT REPEAT 3740 3776
 FT REPEAT 3777 3813
 FT REPEAT 3814 3850
 FT REPEAT 3851 3887
 FT REPEAT 3888 3924
 FT REPEAT 3925 3961
 FT REPEAT 3962 3998
 FT REPEAT 3999 4035
 FT REPEAT 4036 4072
 FT REPEAT 4073 4109
 FT REPEAT 4110 4146
 FT REPEAT 4147 4183
 FT REPEAT 4184 4220
 FT REPEAT 4221 4257
 FT REPEAT 4258 4294
 FT REPEAT 4295 4331
 FT REPEAT 4332 4368
 FT REPEAT 4369 4405
 FT REPEAT 4406 4442
 FT REPEAT 4443 4479
 FT REPEAT 4480 4516
 FT REPEAT 4517 4553
 FT REPEAT 4554 4590
 FT REPEAT 4591 4627
 FT REPEAT 4628 4664
 FT REPEAT 4665 4701
 FT REPEAT 4702 4738
 FT REPEAT 4739 4775
 FT REPEAT 4776 4812
 FT REPEAT 4813 4849
 FT REPEAT 4850 4886
 FT REPEAT 4887 4923
 FT REPEAT 4924 4960
 FT REPEAT 4961 4997
 FT REPEAT 4998 5034
 FT REPEAT 5035 5071
 FT REPEAT 5072 5108
 FT REPEAT 5109 5145
 FT REPEAT 5146 5182
 FT REPEAT 5183 5219
 FT REPEAT 5220 5256
 FT REPEAT 5257 5293
 FT REPEAT 5294 5330
 FT REPEAT 5331 5367
 FT REPEAT 5368 5404
 FT REPEAT 5405 5441
 FT REPEAT 5442 5478
 FT REPEAT 5479 5515
 FT REPEAT 5516 5552
 FT REPEAT 5553 5589
 FT REPEAT 5590 5626
 FT REPEAT 5627 5663
 FT REPEAT 5664 5700
 FT REPEAT 5701 5737
 FT REPEAT 5738 5774
 FT REPEAT 5775 5811
 FT REPEAT 5812 5848
 FT REPEAT 5849 5885
 FT REPEAT 5886 5922
 FT REPEAT 5923 5959
 FT REPEAT 5960 5996
 FT REPEAT 5997 6033
 FT REPEAT 6034 6070
 FT REPEAT 6071 6107
 FT REPEAT 6108 6144
 FT REPEAT 6145 6181
 FT REPEAT 6182 6218
 FT REPEAT 6219 6255
 FT REPEAT 6256 6292
 FT REPEAT 6293 6329
 FT REPEAT 6330 6366
 FT REPEAT 6367 6403
 FT REPEAT 6404 6440
 FT REPEAT 6441 6477
 FT REPEAT 6478 6514
 FT REPEAT 6515 6551
 FT REPEAT 6552 6588
 FT REPEAT 6589 6625
 FT REPEAT 6626 6662
 FT REPEAT 6663 6699
 FT REPEAT 6700 6736
 FT REPEAT 6737 6773
 FT REPEAT 6774 6810
 FT REPEAT 6811 6847
 FT REPEAT 6848 6884
 FT REPEAT 6885 6921
 FT REPEAT 6922 6958
 FT REPEAT 6959 6995
 FT REPEAT 6996 7032
 FT REPEAT 7033 7069
 FT REPEAT 7070 7106
 FT REPEAT 7107 7143
 FT REPEAT 7144 7180
 FT REPEAT 7181 7217
 FT REPEAT 7218 7254
 FT REPEAT 7255 7291
 FT REPEAT 7292 7328
 FT REPEAT 7329 7365
 FT REPEAT 7366 7402
 FT REPEAT 7403 7439
 FT REPEAT 7440 7476
 FT REPEAT 7477 7513
 FT REPEAT 7514 7550
 FT REPEAT 7551 7587
 FT REPEAT 7588 7624
 FT REPEAT 7625 7661
 FT REPEAT 7662 7698
 FT REPEAT 7699 7735
 FT REPEAT 7736 7772
 FT REPEAT 7773 7809
 FT REPEAT 7810 7846
 FT REPEAT 7847 7883
 FT REPEAT 7884 7920
 FT REPEAT 7921 7957
 FT REPEAT 7958 7994
 FT REPEAT 7995 8031
 FT REPEAT 8032 8068
 FT REPEAT 8069 8105
 FT REPEAT 8106 8142
 FT REPEAT 8143 8179
 FT REPEAT 8180 8216
 FT REPEAT 8217 8253
 FT REPEAT 8254 8290
 FT REPEAT 8291 8327
 FT REPEAT 8328 8364
 FT REPEAT 8365 8401
 FT REPEAT 8402 8438
 FT REPEAT 8439 8475
 FT REPEAT 8476 8512
 FT REPEAT 8513 8549
 FT REPEAT 8550 8586
 FT REPEAT 8587 8623
 FT REPEAT 8624 8660
 FT REPEAT 8661 8697
 FT REPEAT 8698 8734
 FT REPEAT 8735 8771
 FT REPEAT 8772 8808
 FT REPEAT 8809 8845
 FT REPEAT 8846 8882
 FT REPEAT 8883 8919
 FT REPEAT 8920 8956
 FT REPEAT 8957 8993
 FT REPEAT 8994 9030
 FT REPEAT 9031 9067
 FT REPEAT 9068 9104
 FT REPEAT 9105 9141
 FT REPEAT 9142 9178
 FT REPEAT 9179 9215
 FT REPEAT 9216 9252
 FT REPEAT 9253 9289
 FT REPEAT 9290 9326
 FT REPEAT 9327 9363
 FT REPEAT 9364 9400
 FT REPEAT 9401 9437
 FT REPEAT 9438 9474
 FT REPEAT 9475 9511
 FT REPEAT 9512 9548
 FT REPEAT 9549 9585
 FT REPEAT 9586 9622
 FT REPEAT 9623 9659
 FT REPEAT 9660 9696
 FT REPEAT 9697 9733
 FT REPEAT 9734 9770
 FT REPEAT 9771 9807
 FT REPEAT 9808 9844
 FT REPEAT 9845 9881
 FT REPEAT 9882 9918
 FT REPEAT 9919 9955
 FT REPEAT 9956 9992
 FT REPEAT 9993 10029
 FT REPEAT 10030 10066
 FT REPEAT 10067 10103
 FT REPEAT 10104 10140
 FT REPEAT 10141 10177
 FT REPEAT 10178 10214
 FT REPEAT 10215 10251
 FT REPEAT 10252 10288
 FT REPEAT 10289 10325
 FT REPEAT 10326 10362
 FT REPEAT 10363 10399
 FT REPEAT 10400 10436
 FT REPEAT 10437 10473
 FT REPEAT 10474 10510
 FT REPEAT 10511 10547
 FT REPEAT 10548 10584
 FT REPEAT 10585 10621
 FT REPEAT 10622 10658
 FT REPEAT 10659 10695
 FT REPEAT 10696 10732
 FT REPEAT 10733 10769
 FT REPEAT 10770 10806
 FT REPEAT 10807 10843
 FT REPEAT 10844 10880
 FT REPEAT 10881 10917
 FT REPEAT 10918 10954
 FT REPEAT 10955 10991
 FT REPEAT 10992 11028
 FT REPEAT 11029 11065
 FT REPEAT 11066 11102
 FT REPEAT 11103 11139
 FT REPEAT 11140 11176
 FT REPEAT 11177 11213
 FT REPEAT 11214 11250
 FT REPEAT 11251 11287
 FT REPEAT 11288 11324
 FT REPEAT 11325 11361
 FT REPEAT 11362 11398
 FT REPEAT 11399 11435
 FT REPEAT 11436 11472
 FT REPEAT 11473 11509
 FT REPEAT 11510 11546
 FT REPEAT 11547 11583
 FT REPEAT 11584 11620
 FT REPEAT 11621 11657
 FT REPEAT 11658 11694
 FT REPEAT 11695 11731
 FT REPEAT 11732 11768
 FT REPEAT 11769 11805
 FT REPEAT 11806 11842
 FT REPEAT 11843 11879
 FT REPEAT 11880 11916
 FT REPEAT 11917 11953
 FT REPEAT 11954 11990
 FT REPEAT 11991 12027
 FT REPEAT 12028 12064
 FT REPEAT 12065 12101
 FT REPEAT 12102 12138
 FT REPEAT 12139 12175
 FT REPEAT 12176 12212
 FT REPEAT 12213 12249
 FT REPEAT 12250 12286
 FT REPEAT 12287 12323
 FT REPEAT 12324 12360
 FT REPEAT 12361 12397
 FT REPEAT 12398 12434
 FT REPEAT 12435 12471
 FT REPEAT 12472 12508
 FT REPEAT 12509 12545
 FT REPEAT 12546 12582
 FT REPEAT 12583 12619
 FT REPEAT 12620 12656
 FT REPEAT 12657 12693
 FT REPEAT 12694 12730
 FT REPEAT 12731 12767
 FT REPEAT 12768 12804
 FT REPEAT 12805 12841
 FT REPEAT 12842 12878
 FT REPEAT 12879 12915
 FT REPEAT 12916 12952
 FT REPEAT 12953 12989
 FT REPEAT 12990 13026
 FT REPEAT 13027 13063
 FT REPEAT 13064 13100
 FT REPEAT 13101 13137
 FT REPEAT 13138 13174
 FT REPEAT 13175 13211
 FT REPEAT 13212 13248
 FT REPEAT 13249 13285
 FT REPEAT 13286 13322
 FT REPEAT 13323 13359
 FT REPEAT 13360 13396
 FT REPEAT 13397 13433
 FT REPEAT 13434 13470
 FT REPEAT 13471 13507
 FT REPEAT 13508 13544
 FT REPEAT 13545 13581
 FT REPEAT 13582 13618
 FT REPEAT 13619 13655
 FT REPEAT 13656 13692
 FT REPEAT 13693 13729
 FT REPEAT 13730 13766
 FT REPEAT 13767 13803
 FT REPEAT 13804 13840
 FT REPEAT 13841 13877
 FT REPEAT 13878 13914
 FT REPEAT 13915 13951
 FT REPEAT 13952 13988
 FT REPEAT 13989 14025
 FT REPEAT 14026 14062
 FT REPEAT 14063 14099
 FT REPEAT 14100 14136
 FT REPEAT 14137 14173
 FT REPEAT 14174 14210
 FT REPEAT 14211 14247
 FT REPEAT 14248 14284
 FT REPEAT 14285 14321
 FT REPEAT 14322 14358
 FT REPEAT 14359 14395
 FT REPEAT 14396 14432
 FT REPEAT 14433 14469
 FT REPEAT 14470 14506
 FT REPEAT 14507 14543
 FT REPEAT 14544 14580
 FT REPEAT 14581 14617
 FT REPEAT 14618 14654
 FT REPEAT 14655 14691
 FT REPEAT 14692 14728
 FT REPEAT 14729 14765
 FT REPEAT 14766 14802
 FT REPEAT 14803 14839
 FT REPEAT 14840 14876
 FT REPEAT 14877 14912
 FT REPEAT 14913 14949
 FT REPEAT 14950 14986
 FT REPEAT 14987 15023
 FT REPEAT 15024 15060
 FT REPEAT 15061 15097
 FT REPEAT 15098 15133
 FT REPEAT 15134 15170
 FT REPEAT 15171 15207
 FT REPEAT 15208 15244
 FT REPEAT 15245 15281
 FT REPEAT 15282 15318
 FT REPEAT 15319 15355
 FT REPEAT 15356 15392
 FT REPEAT 15393 15429
 FT REPEAT 15430 15466
 FT REPEAT 15467 15503
 FT REPEAT 15504 15540
 FT REPEAT 15541 15577
 FT REPEAT 15578 15614
 FT REPEAT 15615 15651
 FT REPEAT 15652 15688
 FT REPEAT 15689 15725
 FT REPEAT 15726 15762
 FT REPEAT 15763 15799
 FT REPEAT 15800 15836
 FT REPEAT 15837 15873
 FT REPEAT 15874 15910
 FT REPEAT 15911 15947
 FT REPEAT 15948 15984
 FT REPEAT 15985 16021
 FT REPEAT 16022 16058
 FT REPEAT 16059 16095
 FT REPEAT 16096 16132
 FT REPEAT 16133 16169
 FT REPEAT 16170 16206
 FT REPEAT 16207 16243
 FT REPEAT 16244 16280
 FT REPEAT 16281 16317
 FT REPEAT 16318 16354
 FT REPEAT 16355 16391
 FT REPEAT 16392 16428
 FT REPEAT 16429 16465
 FT REPEAT 16466 16502
 FT REPEAT 16503 16539
 FT REPEAT 16540 16576
 FT REPEAT 16577 16613
 FT REPEAT 16614 16650
 FT REPEAT 16651 16687
 FT REPEAT 16688 16724
 FT REPEAT 16725 16761
 FT REPEAT 16762 16798
 FT REPEAT 16799 16835
 FT REPEAT 16836 16872
 FT REPEAT 16873 16909
 FT REPEAT 16910 16946
 FT REPEAT 16947 16983
 FT REPEAT 16984 17020
 FT REPEAT 17021 17057
 FT REPEAT 17058 17094
 FT REPEAT 17095 17131
 FT REPEAT 17132 17168
 FT REPEAT 17169 17205
 FT REPEAT 17206 17242
 FT REPEAT 17243 17279
 FT REPEAT 17280 17316
 FT REPEAT 17317 17353
 FT REPEAT 17354 17390
 FT REPEAT 17391 17427
 FT REPEAT 17428 17464
 FT REPEAT 17465 17501
 FT REPEAT 17502 17538
 FT REPEAT 17539 17575
 FT REPEAT 17576 17612
 FT REPEAT 17613 17649
 FT REPEAT 17650 17686
 FT REPEAT 17687 17723
 FT REPEAT 17724 17760
 FT REPEAT 17761 17797
 FT REPEAT 17798 17834
 FT REPEAT 17835 17871
 FT REPEAT 17872 17908
 FT REPEAT 17909 17945
 FT REPEAT 17946 17982
 FT REPEAT 17983 18019
 FT REPEAT 18020 18056
 FT REPEAT 18057 18093
 FT REPEAT 18094 18130
 FT REPEAT 18131 18167
 FT REPEAT 18168 18204
 FT REPEAT 18205 18241
 FT REPEAT 18242 18278
 FT REPEAT 18279 18315
 FT REPEAT 18316 18352
 FT REPEAT 18353 18389
 FT REPEAT 18390 18426
 FT REPEAT 18427 18463
 FT REPEAT 18464 18500
 FT REPEAT 18501 18537
 FT REPEAT 18538 18574
 FT REPEAT 18575 18611
 FT REPEAT 18612 18648
 FT REPEAT 18649 18685
 FT REPEAT 18686 18722
 FT REPEAT 18723 18759
 FT REPEAT 18760 18796
 FT REPEAT 18797 18833
 FT REPEAT 18834 18870
 FT REPEAT 18871 18907
 FT REPEAT 18908 18944
 FT REPEAT 18945 18980
 FT REPEAT 18981 19017
 FT REPEAT 19018 19054
 FT REPEAT 19055 19090
 FT REPEAT 19091 19127
 FT REPEAT 19128 19164
 FT REPEAT 19165 19201
 FT REPEAT 19202 19238
 FT REPEAT 19239 19275
 FT REPEAT 19276 19312
 FT REPEAT 19313 19349
 FT REPEAT 19350 19386
 FT REPEAT 19387 19423
 FT REPEAT 19424 19460
 FT REPEAT 19461 19497
 FT REPEAT 19498 19534
 FT REPEAT 19535 19571
 FT REPEAT 19572 19608
 FT REPEAT 19609 19645
 FT REPEAT 19646 19682
 FT REPEAT 19683 19719
 FT REPEAT 19720 19756
 FT REPEAT 19757 19793
 FT REPEAT 19794 19830
 FT REPEAT 19831 19867
 FT REPEAT 19868 19904
 FT REPEAT 19905 19941
 FT REPEAT 19942 19978
 FT REPEAT 19979 20015
 FT REPEAT 20016 20052
 FT REPEAT 20053 20089
 FT REPEAT 20090 20126
 FT REPEAT 20127 20163
 FT REPEAT 20164 20200
 FT REPEAT 20201 20237
 FT REPEAT 20238 20274
 FT REPEAT 20275 20311
 FT REPEAT 20312 20348
 FT REPEAT 20349 20385
 FT REPEAT 20386 20422
 FT REPEAT 20423 20459
 FT REPEAT 20460 20496
 FT REPEAT 20497 20533
 FT REPEAT 20534 20570
 FT REPEAT 20571 20607
 FT REPEAT 20608 20644
 FT REPEAT 20645 20681
 FT REPEAT 20682 20718
 FT REPEAT 20719 20755
 FT REPEAT 20756 20792
 FT REPEAT 20793 20829
 FT REPEAT 20830 20866
 FT REPEAT 20867 20903
 FT REPEAT 20904 20940
 FT REPEAT 20941 20977
 FT REPEAT 20978 21014
 FT REPEAT 21015 21051
 FT REPEAT 21052 21088
 FT REPEAT 21089 21125

FT	DMAIN	2265	2276	POLY-GLN (OPA
FT	DISULFID	25	35	BY SIMILARITY
FT	DISULFID	29	45	BY SIMILARITY
FT	DISULFID	47	56	BY SIMILARITY
FT	DISULFID	62	73	BY SIMILARITY
FT	DISULFID	67	96	BY SIMILARITY
FT	DISULFID	88	97	BY SIMILARITY
FT	DISULFID	105	116	BY SIMILARITY
FT	DISULFID	110	126	BY SIMILARITY
FT	DISULFID	128	137	BY SIMILARITY
FT	DISULFID	143	154	BY SIMILARITY
FT	DISULFID	148	163	BY SIMILARITY
FT	DISULFID	165	174	BY SIMILARITY
FT	DISULFID	181	194	BY SIMILARITY
FT	DISULFID	188	203	BY SIMILARITY
FT	DISULFID	205	214	BY SIMILARITY
FT	DISULFID	221	232	BY SIMILARITY
FT	DISULFID	226	242	BY SIMILARITY
FT	DISULFID	244	253	BY SIMILARITY
FT	DISULFID	260	271	BY SIMILARITY
FT	DISULFID	265	280	BY SIMILARITY
FT	DISULFID	282	291	BY SIMILARITY
FT	DISULFID	298	311	BY SIMILARITY
FT	DISULFID	305	320	BY SIMILARITY
FT	DISULFID	322	331	BY SIMILARITY
FT	DISULFID	338	349	BY SIMILARITY
FT	DISULFID	343	358	BY SIMILARITY
FT	DISULFID	360	369	BY SIMILARITY
FT	DISULFID	375	386	BY SIMILARITY
FT	DISULFID	380	397	BY SIMILARITY
FT	DISULFID	399	408	BY SIMILARITY
FT	DISULFID	415	428	BY SIMILARITY
FT	DISULFID	422	437	BY SIMILARITY
FT	DISULFID	439	448	BY SIMILARITY
FT	DISULFID	455	466	BY SIMILARITY
FT	DISULFID	460	475	BY SIMILARITY
FT	DISULFID	477	486	BY SIMILARITY
FT	DISULFID	493	503	BY SIMILARITY
FT	DISULFID	498	512	BY SIMILARITY
FT	DISULFID	514	523	BY SIMILARITY
FT	DISULFID	530	541	BY SIMILARITY
FT	DISULFID	535	550	BY SIMILARITY
FT	DISULFID	552	561	BY SIMILARITY
FT	DISULFID	558	578	BY SIMILARITY
FT	DISULFID	573	587	BY SIMILARITY
FT	DISULFID	589	598	BY SIMILARITY
FT	DISULFID	605	616	BY SIMILARITY
FT	DISULFID	610	625	BY SIMILARITY
FT	DISULFID	627	636	BY SIMILARITY
FT	DISULFID	643	653	BY SIMILARITY
FT	DISULFID	648	662	BY SIMILARITY
FT	DISULFID	664	673	BY SIMILARITY
FT	DISULFID	680	691	BY SIMILARITY
FT	DISULFID	685	700	BY SIMILARITY
FT	DISULFID	702	711	BY SIMILARITY
FT	DISULFID	718	728	BY SIMILARITY
FT	DISULFID	723	737	BY SIMILARITY
FT	DISULFID	739	748	BY SIMILARITY
FT	DISULFID	755	766	BY SIMILARITY
FT	DISULFID	760	775	BY SIMILARITY
FT	DISULFID	777	786	BY SIMILARITY
FT	DISULFID	793	804	BY SIMILARITY
FT	DISULFID	798	813	BY SIMILARITY
FT	DISULFID	815	824	BY SIMILARITY
FT	DISULFID	831	842	BY SIMILARITY
FT	DISULFID	836	853	BY SIMILARITY
FT	DISULFID	855	864	BY SIMILARITY
FT	DISULFID	871	882	BY SIMILARITY
FT	DISULFID	876	891	BY SIMILARITY
FT	DISULFID	893	902	BY SIMILARITY
FT	DISULFID	909	920	BY SIMILARITY
FT	DISULFID	914	929	BY SIMILARITY
FT	DISULFID	931	940	BY SIMILARITY

FT	DISULFID	947	958	BY SIMILARITY.
FT	DISULFID	952	967	BY SIMILARITY.
FT	DISULFID	969	978	BY SIMILARITY.
FT	DISULFID	1023	1034	BY SIMILARITY.
FT	DISULFID	1028	1043	BY SIMILARITY.
FT	DISULFID	1045	1054	BY SIMILARITY.
Query Match				
Best Local Similarity		10.8%;	Score 332;	DB 1; Length 2437;
Matches 95;		Conservative 26;	Mismatches 99;	Indels 112; Gaps
Oy	39	GVCHY-----GTRKACCYGM-----RRNSKGVCEATCEBG-----CKFGE-----	73	
Db	309	GTCHNTIGGFHCFCVYNGWTDGDCSENIDDCASAAKSHGATCHRDVASFCECHGRTGLL	368	
Oy	74	-----CY-----GPKKCKCFPCYTKTQSOPDYNEGKMPRCOH--RC	109	
Db	369	CHLDADICISNPOKSGNSCDTNPVSGKAICCTCPPEYTSACNODIDECISLANCEHGRC	428	
Oy	110	VNTHGSYKCFELSGH-----MLMPDTC-----V	133	
Db	429	INTNGSFQCKGLQGYEGPRCMBVNECKSNPCQNDATCLQIGGFHCICMPGEYGEVFCQI	488	
Oy	134	NSRRCATMNC-QYSCEDTEESPQCLCPSSGLRLAPNGRDC-LDIDECASKATVLCPTNRRK	191	
Db	489	NSDDCASQPCPLNGRCIDKINSFHCCKPGR-----SSGLQYDVNDEKAS--TPCKNGKAC	541	
Oy	192	VNTHGSYKCKCHIEFELQYISGRFDCIDINECTMDSHTCSHHANCFPTQSGFKCKQGY	251	
Db	542	TDGPRKRYCECTPPF-----SGHCELDINDCA--SSPC-HYGCRCRGVASFCTDCRPG	593	
Oy	252	KGNGLRCSAIPENSKEYLRAP-----GTIKDR	279	
Db	594	TGR-----LCETNINICELSLQPCRNNGTCODR	619	
RESULT 15				
MTN2_HUMAN	STANDARD:	PRT:	956	AA.
ID	MTN2_HUMAN			
AC	000339; Q9NSZ1;			
DT	30-MAY-2000 (Ref. 39, Created)			
DT	16-OCT-2001 (Ref. 40, Last sequence update)			
DT	16-OCT-2001 (Ref. 40, Last annotation update)			
DE	Matrillin-2 precursor.			
CN	MATN2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).			
RX	PubMed-11124542;			
RA	Muratoglu S., Krysan K., Balazs M., Sheng H., Zakany R., Modis L.,			
RT	Kiss I., Deak F.;			
RT	"Primary structure of human matrillin-2, chromosome location of the			
RT	MATN2 gene and conservation of an AT-AC intron in matrillin genes.";			
RL	Cytogenet. Cell Genet. 90:323-327(2000).			
RN	[2]			
RP	SEQUENCE OF 644-956 FROM N.A.			
RX	MEUJINE-97238863; PubMed=9083061;			
RA	Deak F., Plecha D., Bachrati C., Paulsson M., Kiss I.;			
RT	"Primary structure and expression of matrillin-2, the closest relative			
RT	of cartilage matrix protein within the von Willebrand factor type A-			
RL	like module superfamily".			
J.	J. Biol. Chem. 272:9268-9274(1997).			
RN	[3]			
RP	SEQUENCE OF 244-956 FROM N.A.			
RC	TISSUE=Testis;			
RA	Dusterhoef A., Lauber J., Mewes H.-W., Gassenhuber J., Wiemann S.;			
CC	Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.			
CC	-1- FUNCTION: INVOLVED IN MATRIX ASSEMBLY (BY SIMILARITY).			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A			

CC SHORT FORM, ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: CONTAINS 10 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 WMFA DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U69263; AAC51260.2; -
DR EMBL: AL137638; CAB70853.1; ALT_INIT.
DR HSSP: P05099; 1A05.
DR MIM: 602108; -
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR002035; WMFA.
DR Pfam: PF00008; EGF; 8.
DR Pfam: PF00092; vwa; 2.
DR PRINTS: PR00453; WMPADOMAIN.
DR SMART: SM00181; EGF; 10.
DR SMART: SM00179; EGF_CA; 10.
DR SMART: SM00327; VMA; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; 9.
DR PROSITE: PS01186; EGF_2; 9.
DR PROSITE: PS50234; WMFA; 2.
DR EGF-like domain; Signal; Glycoprotein; Repeat; Coiled coil;
KW Alternative splicing.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 57 956 MATRILIN-2.
FT DOMAIN 24 232 WMFA 1.
FT DOMAIN 238 278 EGF-Like 1.
FT DOMAIN 279 319 EGF-Like 2.
FT DOMAIN 320 360 EGF-Like 3.
FT DOMAIN 361 401 EGF-Like 4.
FT DOMAIN 402 442 EGF-Like 5.
FT DOMAIN 443 483 EGF-Like 6.
FT DOMAIN 484 524 EGF-Like 7.
FT DOMAIN 525 565 EGF-Like 8.
FT DOMAIN 566 606 EGF-Like 9.
FT DOMAIN 607 647 EGF-Like 10.
FT DOMAIN 655 830 WMFA 2.
FT DOMAIN 917 955 COILED COIL (POTENTIAL).
FT DISULFID 242 253 BY SIMILARITY.
FT DISULFID 249 262 BY SIMILARITY.
FT DISULFID 264 277 BY SIMILARITY.
FT DISULFID 283 294 BY SIMILARITY.
FT DISULFID 290 303 BY SIMILARITY.
FT DISULFID 305 318 BY SIMILARITY.
FT DISULFID 324 335 BY SIMILARITY.
FT DISULFID 331 344 BY SIMILARITY.
FT DISULFID 346 359 BY SIMILARITY.
FT DISULFID 365 376 BY SIMILARITY.
FT DISULFID 372 385 BY SIMILARITY.
FT DISULFID 387 400 BY SIMILARITY.
FT DISULFID 406 417 BY SIMILARITY.
FT DISULFID 413 426 BY SIMILARITY.
FT DISULFID 428 441 BY SIMILARITY.
FT DISULFID 447 458 BY SIMILARITY.
FT DISULFID 467 482 BY SIMILARITY.
FT DISULFID 488 499 BY SIMILARITY.
FT DISULFID 495 508 BY SIMILARITY.
FT DISULFID 510 523 BY SIMILARITY.
FT DISULFID 529 540 BY SIMILARITY.
FT DISULFID 536 549 BY SIMILARITY.
FT DISULFID 551 564 BY SIMILARITY.
FT DISULFID 570 581 BY SIMILARITY.
FT DISULFID 577 590 BY SIMILARITY.
FT DISULFID 592 605 BY SIMILARITY.

FT DISULFID 611 622 BY SIMILARITY.
FT DISULFID 618 631 BY SIMILARITY.
FT DISULFID 633 646 BY SIMILARITY.
FT CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 890 890 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 861 879 MISSING (IN SHORT ISOFORM).
FT CONFLICT 356 356 K -> E (IN REF. 3).
FT CONFLICT 594 594 V -> E (IN REF. 3).
FT CONFLICT 644 644 R -> G (IN REF. 2).
FT CONFLICT 755 755 F -> L (IN REF. 3).
FT CONFLICT 935 935 L -> F (IN REF. 2).
SQ SEQUENCE 956 AA; 106840 MW; 826B7F347178FC80 CRC64;
Query Match 10.8%; Score 331; DB 1; Length 956;
Best Local Similarity 23.8%; Pred. No. 7.3e-16;
Matches 129; Conservative 64; Mismatches 206; Indels 144; Gaps 24;
QY 41 CHYGTKLACCYGWRNR-SKVCE-----ATCEPGKFGCEVGPNR---CRCEPGYT----- 87
DB 385 CH-----CLKGFTLNPDKTKTCRINYNALNKGCEH-ECVMMESYCRCHRGYTLDPN 437
QY 88 GKTCSQDVNECGMKPRPCOHRCVNTHGSKYKFCFLSGHMLMPD-ATCVNSRTCAMIN--CQ 144
DB 438 GKTCSR-VDHCAQODHGEQCLNTEDSEFVCCSGEFLINEDLTKTSRYDYCLLSDHGE 496
QY 145 YSCDTEEGPOCLCPSSGRLAPNRDCLIDECASGVICYGN----- 188
DB 497 YSCVMNDSFACQCP-EGHVLNRSDEKTCALDKSCALGDHGCSSVSEDSFVCCFEGY 555
QY 189 -----RR-----CVMTFGSYCKCHIGFELQYISGRYCDIDINECTM 225
DB 556 ILREDGKTCRRDVCQALDHGEHCIVNSDSDTYECLVGFILADGR--CRRDVCKS 613
QY 226 DSHTSHANPCNTGSCRKCKQGY--KGNLRCSALPENSVKYLRAAFTIKRIKL 283
DB 614 THNGEH--ICVNNNSYICKCSGEFVLAEDRCKCKTEGPIDLVFYIDGS-----KS 665
QY 284 LAHKSMMKKAKIKNVTEPRTTP--TPKYNLOPN---YEELVSRGNSHGKKGNEEM 339
DB 666 LGEENFEVVKOPVYTIIDSLTISPKAARVGLQYSTVHTETTLNPN-----AKDM 718
QY 340 KEGLEDEKREEK-----ALKNDIERSLRGDVFFPKYNEAGEFGLILVORKA----- 386
DB 719 KKAVAHMKYMGKMTGLALKHMFERSFTGEGARPFSTRVRAAIVFTDGRAQDDVSEW 778
QY 387 -----LNSKLEHKLNLISVDCS-----FNHGICDMK 412
DB 779 ASKAKANGITMYAVGVGAIEEELQEIASEPTNKHILFYAEDPSTMDIEISEKLKICICAL 838
QY 413 QDREDFDMPADRONAIGFYMAVPALAGHKHDKRLKLPLDQPOSFCLLPYRLAG 472
DB 839 EDSGROD-SPAGE-----LPKTYOQPIESEPTVINIODLLSCSNFVQHRHLEE 888
QY 473 DKV 475
DB 889 DNL 891

Search completed: October 17, 2002, 16:00:52
Job time : 27 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2002, 15:54:18 : Search time 31 Seconds

(Without alignments)
3086.006 Million cell updates/sec

Title: US-09-687-860-24
Perfect score: 3060
Sequence: 1 MFLPWSIALPLLLSWAGGF.....VDGVLYSGLCPPSLLSVDD 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SPREMBL_19:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mhc:*
9: sp.organelle:*
10: sp.phage:*
11: sp.plant:*
12: sp.protist:*
13: sp.virus:*
14: sp.vertibrate:*
15: sp.unclassified:*
16: sp.virus:*
17: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	ID	Description
1	3055	99.8	553	4	Q9NZL7	Q9NZL7 homo sapien
2	3049.5	99.7	554	4	Q9NZL7	Q9NZL7 homo sapien
3	3016	98.6	558	4	Q9NF67	Q9NF67 homo sapien
4	2401.5	78.5	550	11	Q9JZT5	Q9JZT5 mus musculu
5	1095	35.8	561	11	Q9JZT5	Q9JZT5 mus musculu
6	1081.5	35.3	578	11	Q9JZT5	Q9JZT5 mus musculu
7	1069.5	35.0	592	11	Q9JZT5	Q9JZT5 mus musculu
8	1056	34.5	592	11	Q9JZT5	Q9JZT5 mus musculu
9	408.5	13.3	2906	11	Q9JZT5	Q9JZT5 mus musculu
10	405	13.2	2809	4	Q96JF8	Q96JF8 homo sapien
11	386	12.6	708	13	P87363	P87363 gallus gall
12	385.5	12.6	2872	11	Q9WU8	Q9WU8 ratu mus norv
13	383.5	12.5	3857	11	Q9WU8	Q9WU8 mus musculu
14	382	12.5	1174	11	Q9WU8	Q9WU8 mus musculu
15	374.5	12.2	2189	5	Q9B105	Q9B105 elimeria ten
16	370	12.1	608	11	Q9DBE2	Q9DBE2 mus musculu

17	369	12.1	937	5	Q9BLJ1	Q9BLJ1 ciona intes
18	365.5	11.9	1574	11	Q9BZ81	Q9BZ81 ratu mus norv
19	362.5	11.8	576	4	Q9Y3V7	Q9Y3V7 homo sapien
20	362.5	11.8	1511	4	Q75412	Q75412 homo sapien
21	361.5	11.8	1587	4	Q00508	Q00508 homo sapien
22	354.5	11.6	955	4	Q96DN2	Q96DN2 homo sapien
23	354.5	11.6	1833	11	Q08999	Q08999 mus musculu
24	352	11.5	704	13	Q73774	Q73774 gallus gall
25	349	11.4	2673	4	Q96SC3	Q96SC3 homo sapien
26	348	11.4	685	11	Q922K8	Q922K8 mus musculu
27	346	11.3	495	4	Q9HBQ5	Q9HBQ5 homo sapien
28	340.5	11.1	1664	5	Q9TVQ2	Q9TVQ2 caenorhabdi
29	339.5	11.1	741	4	Q96K89	Q96K89 homo sapien
30	339	11.1	956	11	Q99K64	Q99K64 mus musculu
31	339	11.1	1764	11	Q35806	Q35806 ratu mus norv
32	338	11.0	528	11	Q9CXD8	Q9CXD8 mus musculu
33	338	11.0	5636	4	Q96RW7	Q96RW7 homo sapien
34	334	10.9	517	4	Q9NP01	Q9NP01 homo sapien
35	333	10.9	2447	13	Q13149	Q13149 fuqu rubrip
36	333	10.9	2653	5	Q25253	Q25253 lucilia cup
37	332	10.8	999	4	Q9N036	Q9N036 homo sapien
38	331	10.8	589	5	Q9TZ51	Q9TZ51 caenorhabdi
39	331	10.8	689	5	Q95N23	Q95N23 caenorhabdi
40	330.5	10.8	961	11	Q9EQC6	Q9EQC6 mus musculu
41	330	10.8	798	5	Q18026	Q18026 caenorhabdi
42	329.5	10.8	681	13	Q42182	Q42182 brachydanio
43	329.5	10.8	2319	11	Q9R172	Q9R172 ratu mus norv
44	328	10.7	1062	11	Q60789	Q60789 mus musculu
45	328	10.7	1713	11	Q88349	Q88349 mus musculu

ALIGNMENTS

RESULT 1
ID Q9NZL7 PRELIMINARY; PRT; 553 AA.
AC Q9NZL7
DT 01-OCR-2000 (TREMUREL. 15, Created)
DT 01-OCR-2000 (TREMUREL. 15, Last sequence update)
DT 01-DEC-2001 (TREMUREL. 19, Last annotation update)
DE EPIDERMAL GROWTH FACTOR REPEAT CONTAINING PROTEIN.
GN EGF6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20079166; PubMed=10610727;
RA Yeung G., Mulero J.J., Bernitsen R.P., Loeb D.B., Drmanac R., Ford J.E.;
RT "Cloning of a novel epidermal growth factor repeat containing gene
RT EGF6: expressed in tumor and fetal tissues.";
RL Genomics 62:304-307(1999).
CC -1 SIMILARITY: CONTAINS 1 MAM DOMAIN.
DR EMBL: AF186084; AAF27812.1; -;
DR HSSP: P00736; IAPQ.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR000998; MAM.
DR Pfam: PF00629; MAM; 1.
DR SMART: SM00179; EGF_CA; 3.
DR SMART: SM00001; EGF_like; 2.
DR SMART: SM00137; MAM; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; 3.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01187; EGF_CA; 3.
DR PROSITE: PS50060; MAM_2; 1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 553 AA; 61314 MW; 2F55F167857DE50 CRC64;

Query Match 99.8%; Score 3055; DB 4; Length 553;
 Best Local Similarity 99.8%; Pred. No. 2,1e-240;
 Matches 552; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPLPWSLALPLLSSVAGFGGMAASARHHGLASARQPGVCHYGTKLACCGWRNRSKV 60
 DB 1 MPLPWSLALPLLSSVAGFGGMAASARHHGLASARQPGVCHYGTKLACCGWRNRSKV 60
 QY 61 CEATCEPGCKGEGCVGPKRCRCPFGYTGKTSQDVNECGMKRPRQOHCVNTHGSKFC 120
 DB 61 CEATCEPGCKGEGCVGPKRCRCPFGYTGKTSQDVNECGMKRPRQOHCVNTHGSKFC 120
 QY 121 LSGHMLMPDATCVNSRTCAMINCQYSCDETEBGPCLCPSSGLRLAPNGRDLIDECAS 180
 DB 121 LSGHMLMPDATCVNSRTCAMINCQYSCDETEBGPCLCPSSGLRLAPNGRDLIDECAS 180
 QY 181 GKVICPYNRRCVNTFGSYCKCHIGFELQYISGRYDCIDINECTMDSHTCSHHANCFTNQ 240
 DB 181 GKVICPYNRRCVNTFGSYCKCHIGFELQYISGRYDCIDINECTMDSHTCSHHANCFTNQ 240
 QY 241 GSFCKCKOGYKGNGLRCSAIPENSVKEVLRAPGTIKDRIKKLAHKNSMKKAKIKNTV 300
 DB 241 GSFCKCKOGYKGNGLRCSAIPENSVKEVLRAPGTIKDRIKKLAHKNSMKKAKIKNTV 300
 QY 301 PEPTPTPKVNLQPFNYEEIYVSRGNSHGKKGNEEKMEKGLDEKREKALKNDIEER 360
 DB 301 PEPTPTPKVNLQPFNYEEIYVSRGNSHGKKGNEEKMEKGLDEKREKALKNDIEER 360
 QY 361 SLRGVFFPKVNEGEFGLILVQRKALTSKLEHKDLNLSVDCSFNHGICDMKODREDFD 420
 DB 361 SLRGVFFPKVNEGEFGLILVQRKALTSKLEHKDLNLSVDCSFNHGICDMKODREDFD 420
 QY 421 WNPADRONAIGFYMAVPAALAGHKKIDIGRLKLLPDLQPOSNFCLLFYRLAGDVKGLRV 480
 DB 421 WNPADRONAIGFYMAVPAALAGHKKIDIGRLKLLPDLQPOSNFCLLFYRLAGDVKGLRV 480
 QY 481 FVKNSNNALAMEKTTSEDEKWKTKIQLYQGTDAKTSIIFEAERKKGKTGEIADVGVLLV 540
 DB 481 FVKNSNNALAMEKTTSEDEKWKTKIQLYQGTDAKTSIIFEAERKKGKTGEIADVGVLLV 540
 QY 541 SGLCPDLSLSDVD 553
 DB 541 SGLCPDLSLSDVD 553

RESULT 2
 Q9NY67 PRELIMINARY; PRT: 554 AA.

AC Q9NY67; TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL 61.4 KDA PROTEIN.
 GN W80.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RN RP SEQUENCE FROM N.A.
 RC TISSUE=TERATOCARCINOMA/NEURON;
 RA Franco B.;
 RL submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RP [2]
 RP TISSUE=TERATOCARCINOMA/NEURON;
 RC MEDLINE=20241927; PubMed=10777661;
 RA Buchner G., Orfanelli U., Quaderi N., Bassi M.T., Andolfi G.;
 RT "Identification of a new BGF-repeat-containing gene from Xp22:
 RT Acandante for developmental disorders.";
 RL Genomics 65:16-23(2000).
 CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.

DR EMBL; AJ245671; CAB92132.1; -.
 DR HSSP; P00736; IAPQ.
 DR InterPro; IPR000152; Asx_Hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR000998; MAM.
 DR Pfam; PF00008; EGF; 4.
 DR Pfam; PF00629; MAM; 1.
 DR SMART; SM00179; EGF_CA; 3.
 DR SMART; SM00001; EGF_Like; 2.
 DR SMART; SM00137; MAM; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; 3.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS01187; EGF_CA; 3.
 DR PROSITE; PS50060; MAM_2; 1.
 KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation;
 KW Hypothetical protein; Repeat.
 SEQUENCE 554 AA: 6138 MW: D519238F2A604101 CRC64;

Query Match 99.7%; Score 3049.5; DB 4; Length 554;
 Best Local Similarity 99.8%; Pred. No. 6e-240;
 Matches 553; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MPLPWSLALPLLSSVAGFGGMAASARHHGLASARQPGVCHYGTKLACCGWRNRSKV 60
 DB 1 MPLPWSLALPLLSSVAGFGGMAASARHHGLASARQPGVCHYGTKLACCGWRNRSKV 60
 QY 61 CEATCEPGCKGEGCVGPKRCRCPFGYTGKTSQDVNECGMKRPRQOHCVNTHGSKFC 120
 DB 61 CEATCEPGCKGEGCVGPKRCRCPFGYTGKTSQDVNECGMKRPRQOHCVNTHGSKFC 120
 QY 121 LSGHMLMPDATCVNSRTCAMINCQYSCDETEBGPCLCPSSGLRLAPNGRDLIDECAS 180
 DB 121 LSGHMLMPDATCVNSRTCAMINCQYSCDETEBGPCLCPSSGLRLAPNGRDLIDECAS 180
 QY 181 GKVICPYNRRCVNTFGSYCKCHIGFELQYISGRYDCIDINECTMDSHTCSHHANCFTNQ 240
 DB 181 GKVICPYNRRCVNTFGSYCKCHIGFELQYISGRYDCIDINECTMDSHTCSHHANCFTNQ 240
 QY 241 GSFCKCKOGYKGNGLRCSAIPENSVKEVLRAPGTIKDRIKKLAHKNSMKKAKIKNTV 300
 DB 241 GSFCKCKOGYKGNGLRCSAIPENSVKEVLRAPGTIKDRIKKLAHKNSMKKAKIKNTV 300
 QY 301 PEPTPTPKVNLQPFNYEEIYVSRGNSHGKKGNEEKMEKGLDEKREKALKNDIEER 360
 DB 301 PEPTPTPKVNLQPFNYEEIYVSRGNSHGKKGNEEKMEKGLDEKREKALKNDIEER 360
 QY 361 SLRGVFFPKVNEGEFGLILVQRKALTSKLEHK-DLWISVDCSFNHGICDMKODREDF 419
 DB 361 SLRGVFFPKVNEGEFGLILVQRKALTSKLEHKDLNLSVDCSFNHGICDMKODREDF 420
 QY 420 WNPADRONAIGFYMAVPAALAGHKKIDIGRLKLLPDLQPOSNFCLLFYRLAGDVKGLRV 479
 DB 420 WNPADRONAIGFYMAVPAALAGHKKIDIGRLKLLPDLQPOSNFCLLFYRLAGDVKGLRV 480
 QY 481 FVKNSNNALAMEKTTSEDEKWKTKIQLYQGTDAKTSIIFEAERKKGKTGEIADVGVLL 540
 DB 481 FVKNSNNALAMEKTTSEDEKWKTKIQLYQGTDAKTSIIFEAERKKGKTGEIADVGVLL 540
 QY 540 VSGLCPSLSDVD 553
 DB 541 VSGLCPSLSDVD 554

RESULT 3
 Q9UFK6 PRELIMINARY; PRT: 558 AA.

AC Q9UFK6; TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL 61.8 KDA PROTEIN (FRAGMENT).

Fri Oct 18 11:29:56 2002

us-09-687-860-24.rsp

Page 3

```

GN DKEP564P2063.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
OX NCBI_TaxID=9606;
RN
[1]
SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Diesterheft A., Lauber J., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.
DR EMBL; AL117610; CAB56014.1; -.
DR HSSP; P00736; IAP0.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR000998; MAM.
DR Pfam; PF00008; EGF; 4.
DR SMART; SM00179; EGF_CA; 3.
DR SMART; SM00001; EGF_Like; 2.
DR SMART; SM00137; MAM; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01187; EGF_CA; 2.
DR PROSITE; PS50060; MAM_2; 1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation;
KW Hypothetical protein; Repeat.
FT NON_TER
SQ SEQUENCE 558 AA; 61828 MW; AA38D7DCE402BFA3 CRC64;

Query Match 98.6%; Score 3016; DB 4; Length 558;
Best Local Similarity 98.9%; Pred. No. 3.2e-237;
Matches 547; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 MRLPMSIALPRLLSVWAGFGNNAASARHNGGLASAROPGVCHYGTAKLACCYGWRNRSKV 60
DB 6 MRLPMSIALPRLLSVWAGFGNNAASARHNGGLASAROPGVCHYGTAKLACCYGWRNRSKV 65
QY 61 CEATCEPGCKGEGCVGNPKRCRPFYGTGKTSQDVNECGMKPRPCQHRVCVTHGSYKFC 120
DB 66 CEATCEPGCKGEGCVGNPKRCRPFYGTGKTSQDVNECGMKPRPCQHRVCVTHGSYKFC 125
QY 121 LSGHMLPDAVCNSTRCAAMINCOYSCDTEEGPQCLOPSSGLRLAPRGRCLOIDECAS 180
DB 126 LSGHMLPDAVCNSTRCAAMINCOYSCDTEEGPQCLOPSSGLRLAPRGRCLOIDECAS 185
QY 181 GKVICPNRRVCVNTFGSYCKHIGFELQYISGRYDCIDINECTMDSHSCSHANCFNTQ 240
DB 186 GKVICPNRRVCVNTFGSYCKHIGFELQYISGRYDCIDINECTMDSHSCSHANCFNTQ 245
QY 241 GSFCKCKGKGYKGLRCSAIPENSVEVLAPGTIKRIKLLAHKNSMKKKAKIKVNT 300
DB 246 GSFCKCKGKGYKGLRCSAIPENSVEVLAPGTIKRIKLLAHKNSMKKKAKIKVNT 305
QY 301 PEPRTRTPKKNLQPFNVEELVSRGNSHGKKGNEEMKKGLEDEKKEEKALNDIEER 360
DB 306 PEPRTRTPKKNLQPFNVEELVSRGNSHGKKGNEEMKKGLEDEKKEEKALNDIEER 365
QY 361 SLRGDVEFPKVNAGEFGLILVQRKALLSKLEHRDLNLSVDCSFNHCICDKODREDFD 420
DB 366 SLRGDVEFPKVNAGEFGLILVQRKALLSKLEHRDLNLSVDCSFNHCICDKODREDFD 425
QY 421 WNPADRONAIFGYMAVPLAGHKDIGRLKLLPDLQPOSNFCLLFDYRLAGDKVYKRLV 480
DB 426 WNPADRONAIFGYMAVPLAGHKDIGRLKLLPDLQPOSNFCLLFDYRLAGDKVYKRLV 485
QY 481 FVKSNNALAMEKTTSEDEKKTGKIOLYOGTDATKSIIFAEARGKGTGELAVDGYLLV 540
DB 486 FVKSNNALAMEKTTSEDEKKTGKIOLYOGTDATKSIIFAEARGKGTGELAVDGYLLV 545
QY 541 SGLCPDLSLVVD 553

```

```

DB 546 SGLCPDLSLVVD 558
|||||
RESULT 4
ID 09J025 PRELIMINARY; PRT; 550 AA.
AC 09J025;
DT 01-OCT-2000 (TReMBLrel. 15 Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE HYPOTHEICAL 61.5 KDA PROTEIN.
GN EGF6 OR W80.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN
[1]
SEQUENCE FROM N.A.
RA Franco B.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN
[2]
SEQUENCE FROM N.A.
RX MEDLINE=20241927; PubMed=10777661;
RA Buchner G., Orfanelli U., Quaderl N., Bassi M.T., Andolfi G.;
RT "Identification of a new EGF-repeat-containing gene from human Xp22:
RT Acandidate for developmental disorders.";
RL Genomics 65:16-23(2000).
CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.
DR EMBL; AJ245672; CAB92118.1; -.
DR HSSP; P35555; IEMN.
DR MGD; MGI:1858599; Egfl6.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR000998; MAM.
DR Pfam; PF00008; EGF; 4.
DR Pfam; PF00629; MAM; 1.
DR SMART; SM00179; EGF_CA; 3.
DR SMART; SM00001; EGF_Like; 2.
DR SMART; SM00137; MAM; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01187; EGF_CA; 3.
DR PROSITE; PS50060; MAM_2; 1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation;
KW Hypothetical protein; Repeat.
SQ SEQUENCE 550 AA; 61520 MW; DEF936325C9F31B3 CRC64;

Query Match 78.5%; Score 2401.5; DB 11; Length 550;
Best Local Similarity 77.9%; Pred. No. 3.4e-187;
Matches 431; Conservative 48; Mismatches 69; Indels 5; Gaps 4;

QY 1 MRLPMSIALPRLLSVWAGFGNNAASARHNGGLASAROPGVCHYGTAKLACCYGWRNRSKV 60
DB 1 MRLPMSIALPRLLSVWAGFGNNAASARHNGGLASAROPGVCHYGTAKLACCYGWRNRSKV 65
QY 61 CEATCEPGCKGEGCVGNPKRCRPFYGTGKTSQDVNECGMKPRPCQHRVCVTHGSYKFC 120
DB 61 CEATCEPGCKGEGCVGNPKRCRPFYGTGKTSQDVNECGMKPRPCQHRVCVTHGSYKFC 120
QY 121 LSGHMLPDAVCNSTRCAAMINCOYSCDTEEGPQCLOPSSGLRLAPRGRCLOIDECAS 180
DB 121 LSGHMLPDAVCNSTRCAAMINCOYSCDTEEGPQCLOPSSGLRLAPRGRCLOIDECAS 180
QY 181 GKVICPNRRVCVNTFGSYCKHIGFELQYISGRYDCIDINECTMDSHSCSHANCFNTQ 240
DB 181 GKVICPNRRVCVNTFGSYCKHIGFELQYISGRYDCIDINECTMDSHSCSHANCFNTQ 240
QY 241 GSFCKCKGKGYKGLRCSAIPENSVEVLAPGTIKRIKLLAHKNSMKKKAKIKVNT 300
DB 241 GSFCKCKGKGYKGLRCSAIPENSVEVLAPGTIKRIKLLAHKNSMKKKAKIKVNT 300
QY 301 PEPRTRTPKKNLQPFNVEELVSRGNSHGKKGNEEMKKGLEDEKKEEKALNDIEER 360
DB 301 PEPRTRTPKKNLQPFNVEELVSRGNSHGKKGNEEMKKGLEDEKKEEKALNDIEER 360
QY 361 SLRGDVEFPKVNAGEFGLILVQRKALLSKLEHRDLNLSVDCSFNHCICDKODREDFD 420
DB 361 SLRGDVEFPKVNAGEFGLILVQRKALLSKLEHRDLNLSVDCSFNHCICDKODREDFD 420
QY 421 WNPADRONAIFGYMAVPLAGHKDIGRLKLLPDLQPOSNFCLLFDYRLAGDKVYKRLV 480
DB 421 WNPADRONAIFGYMAVPLAGHKDIGRLKLLPDLQPOSNFCLLFDYRLAGDKVYKRLV 480
QY 481 FVKSNNALAMEKTTSEDEKKTGKIOLYOGTDATKSIIFAEARGKGTGELAVDGYLLV 540
DB 481 FVKSNNALAMEKTTSEDEKKTGKIOLYOGTDATKSIIFAEARGKGTGELAVDGYLLV 540
QY 541 SGLCPDLSLVVD 553

```

Oy	301	PEPPTPTPKNLDPENFEYVSGGNSHGKCKNEEKMEGLEDKERREKALKNDE -	359
Oy	301	PEPPTPTPKNLDPENFEYVSGGNSHGKCKNEEKMEGLEDKERREKALKNDE -	359
Db	299	PRPASTRYPKVNL - PYSEEBVSGSRNITDGGKKKEGKREPLEER - GERTLRNEVQE	356
Oy	360	RSLRGDEVFPKVNAGEFGILLVORKALTSKLEHKDNLISVDCSFNHGICDMKODREDF	419
Db	357	RTLREDVPSPKVNEAEDLDLVYVORKELEKTKHKDLNISVDCSFGDGYCDMKODREDDF	416
Oy	420	DWNPADRONALGFVMAVPALGHHKDDIGRKLILPDLPOPSNFCLLPDYRLAGKGVKLR	479
Db	417	DWNPADRONDGYTMAVPALGHHKKNIGRLILPDLNLPPOSNFCLLPDYRLAGDKVKLR	476
Oy	480	VFKVNSNNALAMEKTTSEDEXKWKTKIQLVYGTDATKSIIIEAERGKGTGEIAYDGVLL	539
Db	477	VFKVNSNNALAMEETKKNEDGMRGTKIQLVYGIDTTSVIEAERGKGTGEIAYDGVLL	536
Oy	540	VSGICPDLSLSDV 552	
Db	537	VSGICPDPLSLVE 549	

RESULT 5	
Q91V88	
ID Q91V88	PRELIMINARY;
511188	PRT; 561 AA

DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE POEM (NEPHRONECTIN SHORT ISOFORM).
 GN POEM OR NEPH1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RA MEDLINE=21551216; PubMed=11546798;
 RA Morimura N., Tezuka Y., Watanabe N., Yasuda M., Miyatani S.,
 Horumi N., Tezuka K.;
 RT "Molecular cloning of POEM, A novel adhesion molecule that interacts
 RT with alpha8beta1 integrin.";
 RL J. Biol. Chem. 276:42172-42181(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NIH SWISS; TISSUE=KIDNEY;
 RX MEDLINE=21363579; PubMed=11470831;
 RA Brandenberger R., Schmidt A., Linton J., Wang D., Backus C., Denda S.
 RA Muller U., Ketchard L.F.;
 RT "Identification and characterization of a novel extracellular matrix
 RT protein nephronectin that is associated with Integrin alpha8beta1 in
 RT the embryonic kidney.";
 RL J. Cell Biol. 154:447-458(2001).
 DR EMBL; AB059656; BAB69692.1; -;
 DR EMBL; AY035898; AAK36010.1; -;
 SO SEQUENCE 561 AA; 61490 MW; 69E7ACAA0EE3F506 CRC64;

Query Match	35.8%	Score 1095;	DB 11;	Length 561;
Best Local Similarity	38.0%	Pred. No. 8e-81;		
Matches 210; Conservative	92;	Mismatches 180;	Indels 70;	Gaps 11;

0Y	39	GVAHHGTLACCYGARRRASKVCEATBPGCKBEECYGNKRCRPFYRTKTSQDYNEC	98
		: : : : : : : : : : : : : : : : : : :	
Db	34	GLCRGGRIIDCCWGNARSKWQCPVQOPQCKHBEYGNPKCKHPFAFKTACQDIDNEC	93
0Y	99	GAKPRPCQHRCVNTHGSKYKCFCLSGHMLMPATQTVNSRTAMLCQYSCEDTEBGPQLC	158
		: : : : : : : : : : : : : : : : : : :	
Db	94	GLKPRPCNHRKMNTFFGSKYKCYCLGMYMLLPDGSCSSALSCSMANCOYGCVWNGQVNCOC	153
0Y	159	PSSGRLRLPNCRDCLDIDECASGAVICPCYNRCVNTFASGYKCHIEFELYTSGRDCL	218
		: : : : : : : : : : : : : : : : : : :	
Db	154	PSPGQLAPBDRCTQDIDECATGNVSCRFQACQVNTGSKYKCHTFIDMLYIGGKYOCH	213

```

QY      219  DINETMDSHCNHNANFNQGFCKCKQGYKNGICRCSAIPENSVK-----EVLRAP 2733
Db      214  DDECSLQHQCCSSTARCTINHGSIKCCRCRGITGDDGNCVYIPKVIIEFSGPIHMEERN 2733
QY      274  GTTK-----DRIKKLAAHKSMMKKAK-INKVTPERTPTPTPK 310
Db      274  GTISKGGDGHANRIPDAGSTRWPLKTPYIPVIINRPTSKPTTPRPPTNPPQPTPTPTPP 3333
QY      311  VNLRPNTEEIVSRGNSHGKCKGNEEMKEGLEDEKREKAL-----KN 3555
Db      334  LPTERT-----RTPLPTPTPTPTPTPTTIPATSTTTRVITVDNRI 374
QY      356  DIEERSLRGDVFPPPVNAGEFGLIVORAKLTSLKLEHKDNLNVS-DCSFNHGICOMKOD 414
Db      375  QTDPOKPRGPDVPIRQPPNDLFEIPEIERGVSADEEVKDDPGIILHSCNPDHGLCGWIRE 434
QY      415  REDDEDMNPADRDNAIGFYMAVPPALAGHKDIGHRIKLLPDLQPSNCFLLFYRLAGDK 474
Db      435  KDSLDLHWETA-RDPAQGGYLTVSAAKAPGCAARLVLRLGILHMSDGLCSLFRKVTGLH 4933
QY      475  VGKLRVEYK--NSNNALAMEKTTSEDEKMKTGKIQLYQGTDAIKSTIFAEERCKGTGEI 5522
Db      494  SQTLOVFEYRKHTGHAALMGNGCHG--WROTQITL-RGAD-VKSVTFEKEKRGHTGEI 549
QY      533  ANVDGVLVSGLC 544
Db      550  GUDDVSLKRGRC 561

```

RESULT 6	
Q91ZD3	
ID Q91ZD3	PRELIMINARY; PRT; 578 AA

DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE NEPHRONECTIN LONG ISOFORM.
GN NEPH1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Scurionath; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NH SWISS; TISSUE=KIDNEY;
RX MEDLINE=21363579; PubMed=11470831;
RA Brandenberger R., Schmidt A., Linton J., Wang D., Backus C., Denda S.,
RA Muller U., Reichardt L.F.;
RT "Identification and characterization of a novel extracellular matrix
RT protein nephronectin that is associated with integrin alpha5beta1 in
RT the embryonic kidney *";
RL J. Cell Biol. 154:447-458(2001).
SO EMBL; AY035899; AAK96011.1; -
SQ SEQUENCE 578 AA; 63606 MW; 4757DE54CC47DA4C CRC64;

Query Match	35.3%	Score	1081.5	DB	11	Length	578
Best Local Similarity	37.1%	Pred.	No. 1e-79				
Matches	211	Conservative	92	Mismatches	179	Indels	87
						Gaps	12

[illegible]

```

QY      94 -----DVNEGCKM.PRPQOHRVCVNFHSGYSYKCFCLSGHMLM 127
Db      34 GLCRGRGIDDCMGMARNSWGCQCPVQCPOCKRHECVCYGNPKCKCHPFAKTKTQNDSSFH 93
QY      39 GVCCHGTGLACICGYRRNRNSKVCVATCEPGRCKFECCYGNPKRCFPEYTKGTSQ----- 93
Db      34 GLCRGRGIDDCMGMARNSWGCQCPVQCPOCKRHECVCYGNPKCKCHPFAKTKTQNDSSFH 93
QY      94 -----DVNEGCKM.PRPQOHRVCVNFHSGYSYKCFCLSGHMLM 127
Db      94 PTPLDGSEGPLEPQRPDHQATNPVSRDLNECGLRKPRCKHRCMNTFSGYKCYCLNGYMLL 153
QY      128 PDATCVNSRTCAMINCOYSCEDTEBGPQCLCPSSGLRLAPNGRCLDIDBCASGKVTCPY 187
Db      154 PDGSSCSALSCSMANCOYGCQVNVKGOVRCQCPSPGLDILAPGRCTVDIDECATGRVSCPR 213
QY      188 NRRCVNFVFGSYCKCHIGFELYISGRDCLDIDECMTSDHTSCSHANCFRTQGSFCKC 247
Db      214 FRQCVNTFGSYCKCHHTGFEDLMTYIGGKQCHDIDEGCSGHOQGSYKRCVNIHNSYGCOC 273

```

[illegible]

QY 278 -----DRIKLLAHKNSMKKAK-1KNVTPERTPTPKVNLQPFVVEIV 322
 DB 334 RIPDAGSRMPLKPTPIPTNRTSKPTTRPNPPTPPPPPLPTP----- 386
 QY 323 SRGNSHGCKGKNEKMEGLEDEKREKAL-----KNDIERSLRGDVF 367
 DB 387 -----RTYPLPTPERTPTPTTTPATSTTTRVITVYDNRLOTQPKPRGDVF 434
 QY 368 FPKVNEAFEGFLIVORALSKLEHKDLNTSY--DCSFNHGICDOKOREDFDNPPDR 426
 DB 435 IPRQPTNLFETFEIERKVSADDEYKDDPGILHSCNFDHGLCGITREKSDLHETA-R 493
 QY 427 DNATGYAVAPALAGHKKIDGRLLKLLPDLQPOSNCLLFYRLAGDYKGLRFEVK--N 484
 DB 494 DPAGGQYLTVSAAKAPGKARLVLRGLHMSGDLCSFRHKYVGLSHGTLQVFPVKHG 553
 QY 485 SNNALAEKTTSEDEKMTGKIQLQYGTDAKSTTFEAEKRGKGTGELAVDGVLLVSGIC 544
 DB 554 THGALMGKNGCHG--WRQTOITL-RGAD-VKSVTFKGEKRGHGTGLDVSLSKRGRC 609

RESULT 9

ID 09WU9 PRELIMINARY; PRT; 2906 AA.
 AC 09WU9;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE FIBRILLIN-2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93350231; PubMed=10419698;
 RA Yang Q., Ota K., Tian Y., Kumar A., Wada J., Kashihara N., Wallner E.,
 Kanwar Y.S.;
 RT "Cloning of rat fibrillin-2 cDNA and its role in branching
 morphogenesis of embryonic lung.";
 RL Dev. Biol. 212:229-242(1999).
 DR EMBL; AF135060; AAC34439.1;-
 DR HSSP; P35555; 1EMN.
 DR InterPro; IPR002086; Aldehyde_dehydr.
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR001436; EGF_T1.
 DR InterPro; IPR002212; TB.
 DR Pfam; PF00008; EGF; 46.
 DR Pfam; PF00683; TB; 9.
 DR PRINTS; PRO0010; EGFBLD.
 DR SMART; SM00179; EGF_CA; 42.
 DR SMART; SM00001; EGF_like; 4.
 DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE; PS00010; ASX_HYDROXYL; 43.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE; PS01186; EGF_2; 36.
 DR PROSITE; PS01187; EGF_CA; 43.
 KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
 SQ SEQUENCE 2906 AA; 313371 MW; 9EB64E72704AE58 CRC64;

Query Match 13.3%; Score 408.5; DB 11; Length 2906;
 Best Local Similarity 29.3%; Pred. No. 5.1e-24;
 Matches 98; Conservative 32; Mismatches 92; Indels 113; Gaps 13;

QY 38 PGVCHYGT-----KLACCYG-----RNSKGVCEATCEPG--CKFGEV---GPNKC 80
 DB 1073 PGMCYTGKCRMTIGSFKRCNGSFALDMERNCTDIDECRISPDLCNGMCIVNTPGSEFC 1132
 QY 81 RCFGYTG-----KTCSDVNECGMKPRPCQ-HRCVNTHGSKCKCLSGHMLMP----- 128

DB 1133 ECFEGESGFMMKNC-MDIDECERNPLLCRGGTGCVNTEGSPQCDPLGHELSPSREDCI 1191
 QY 129 -----DATCVNSRTCAMI-----NCQYSCD 149
 DB 1192 DINECSLSDNLCNRKNCVMIMTGYOCSNPGYQATPPDRQSGDIDECMIMNCGCTQCTN 1251
 QY 150 TEEGPOCLPSSGRLAPNRDCLDIDECASGVIC-----PYNRCV----- 192
 DB 1252 SEGSYECSC-SEGYALMPDGRSCADIDECENNPDI CDGQCTNIPGEYRCLCYDGFMSM 1310
 QY 193 -----NTFGSYCKCHIGFELYISGRYCDIDINECTMOSH 228
 DB 1311 DKMTCIDVNECDLNPNICFGECECTKGSFICHQGLYSVK--KATGCTDVDECEIGAH 1368
 QY 229 TCSHANCENFQSEFKCKQKQKGNGLRCSAIP 263
 DB 1369 NCDMAASCLNVPGRFKSCREGVNGNGIKCIDLDE 1403

RESULT 10

ID 096JP8 PRELIMINARY; PRT; 2809 AA.
 AC 096JP8;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE FIBRILLIN3.
 GN KIAA1776.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RX MEDLINE=21245130; PubMed=11347906;
 RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XX.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 8:85-95(2001).
 DR EMBL; AB053450; BAB47408.1;-
 SQ SEQUENCE 2809 AA; 300323 MW; 20C04CC006C016F CRC64;

Query Match 13.2%; Score 405; DB 4; Length 2809;
 Best Local Similarity 33.2%; Pred. No. 9.3e-24;
 Matches 93; Conservative 37; Mismatches 104; Indels 46; Gaps 12;

QY 35 ARQGVCHYGT-----KLACCYGWRNSKG-----VCEATCEPG--CKFGEV---GP 77
 DB 1076 ARDPLLCRGCTNTDGSYKCCPGHEHETFAKGTACEDIDECSLSDGLCPHQCVNAVIGA 1135
 QY 78 NKRCFPGYTGKCSO-----DVNECGMKPRPCQHRVNTGSKYKCFCLSGHMLMPATCVN 134
 DB 1136 FQCSCHAGFQSTPDRQGCYVDINECRVQNGCDYHCINTGSTRCSGGQYSLMPD----- 1190
 QY 135 SRTCAMIN-----C-QYSCDETEEGPOCLPSSGRLAPNRDCLDIDECASGVIC 185
 DB 1191 GRACADVDCEENPRVCDGSHCTNMPGHRCLC-YDGFNATPDMRCYVDDECDLNPHTC 1249
 QY 186 PYNRCAVNTFGSYKCHIGFELYISGRYCDIDINECTMDSHTSHANCAVNTQSGFRC 245
 DB 1250 LHG-DCENTKGSFVCHQGLGYMR--KGATGCSVDDECEVGGHNDSDHASCINTIPGSEFC 1306
 QY 246 KCKQGYKGNGLRCSAIPENSVE-----VLAPGTIK 277
 DB 1307 RCLPGWVGDEFCDDIDECISOEHRCSPRGDCLVNPGSTR 1346

RESULT 11
 ID P87363 PRELIMINARY; PRT; 708 AA.
 AC P87363;

DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE FIBRILIN-1 (FRAGMENT).
 GN FBN1.
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 NCBI_TaxID=9031;
 RN
 RX NCBL_TaxID=9031;
 RP
 RA MEDLINE=20152896; PubMed=10691037;
 RA Zhou G., Price C.E., Rosenquist T.H., Gadsen P.F., Godfrey M.,
 RT "Partial cloning and sequencing of chick fibrillin-1 cDNA."
 RL In Vitro Cell. Dev. Biol. Anim. 36:19-25(2000).
 DR EMBL; U88872; AAB48531.1; -.
 DR HSSP; P07204; 2ADX.
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR002212; TB.
 DR InterPro; IPR001491; Thrbomodulin.
 DR Pfam; PF00683; TB; 2.
 DR PRINTS; PR00907; THRBOMODULN.
 DR SMART; SM00179; EGF_CA; 14.
 DR PROSITE; PS00010; ASX_HYDROXYL; 13.
 DR PROSITE; PS01186; EGF_2; 10.
 DR PROSITE; PS01187; EGF_CA; 13.
 KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
 FT NON_TER
 SQ SEQUENCE 708 AA; 76164 MW; C247271C1DF73361 CRC64;

Query Match 12.6%; Score 386; DB 13; Length 708;
 Best Local Similarity 33.0%; Pred. No. 5.7e-23;
 Matches 89; Conservative 33; Mismatches 88; Indels 60; Gaps 11;

QY 36 ROPGVCHGTACCCYGRNRNKGVCATCEPG-----CKRGE 73
 DB 142 RNPILCRGCTCT-----NTEGSFRCDCPPGHHISPNISACIDINECDLSTNLCRNGH 193
 QY 74 CV---GPNKRCCEPGY---TGKTCSDYNECGMKPRCOHRCVNTGSKFCFLSGHML 127
 DB 194 CVNLIGKQCACNPGYGTADKLHCHIDIDECISINNGCENFTGSEGSCECKGCFALM 253
 QY 128 PDATCVNSRTGAMINCOYSCEDTE---EGPQ-----CLCPSSGLRLAPNGRDCLDI 175
 DB 254 PD-----HRTCTDID---ECEDNPNICDGGCCTNIPGEYRCLC-YDGFMASEDMDKTCVDY 304
 QY 176 DECAAGKICYPYNNRCVMTFGSYCKCHIGFELYISGRYDCIDINECTMDSHTSNHN 235
 DB 305 NECDLHPNIC-LSGTCENPKGSFICHCDMGYSK--KGTGCTDINECHIGAHNCDRHAV 361
 QY 236 CFNTGSEFKCKCKGKGYNGGLRCSAIPENS 265
 DB 362 CTNPGSFKSCSSGWMINGIKCTDLDECS 391

RESULT 12
 Q9WUH8 PRELIMINARY; PRT; 2872 AA.
 AC Q9WUH8:
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE FIBRILIN-1.
 GN FBN1.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RN
 RP SEQUENCE FROM N.A.

RX MEDLINE=99032689; PubMed=9815129;
 RA Kanwar Y.S., Ota K., Yang Q., Kumar A., Wada J., Kashihara N.,
 RA Peterson D.R.,
 RT "Isolation of rat fibrillin-1 cDNA and its relevance in metaphoric
 RT development."
 RT Am. J. Physiol. 275:F710-F723(1998).
 RN
 RP SEQUENCE FROM N.A.
 RA Kanwar Y.S.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF135059; AAD34438.1; -.
 DR HSSP; P35555; 1APJ.
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR002212; TB.
 DR InterPro; IPR000822; ZnF-C2H2.
 DR Pfam; PF00008; EGF; 45.
 DR Pfam; PF00683; TB; 9.
 DR SMART; SM00179; EGF_CA; 41.
 DR SMART; SM00001; EGF_Like; 5.
 DR PROSITE; PS00010; ASX_HYDROXYL; 42.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE; PS01186; EGF_2; 38.
 DR PROSITE; PS01187; EGF_CA; 41.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
 SQ SEQUENCE 2872 AA; 312069 MW; 0C4F7F3B87A80280 CRC64;

Query Match 12.6%; Score 385.5; DB 11; Length 2872;
 Best Local Similarity 28.2%; Pred. No. 3.7e-22;
 Matches 95; Conservative 37; Mismatches 92; Indels 113; Gaps 13;

QY 38 PGVCHGT-----KLACCYGM-----RNSKGVCAATCEPG--CKRGEV---GPNKC 80
 DB 1037 PSLCTHYGKCRNTIGSEFKCSCDGFALDSERRCTDIDECRISPDLCGRQCYNTPEDFEC 1096
 QY 81 RCFPGYTG-----KTCSDOVNECGMKPRCOHR-CVNTGSKFCFLSGHMLPMDT--- 131
 DB 1097 KCDEGESEGFMMKNC-MDIDECQRPDLRCRGICHTNGSTRCECPSSHQSLSPNISACTI 1155
 QY 132 -----CNSRTGAMIN---COYSCED 149
 DB 1156 DINEGELSANLCPHRCVNLIGKYECACNPGYHPTDRFLFCYDIDECISINNGCCEPFCFN 1215
 QY 150 TEEGPOCLCPSSGLRLAPNGRDCLDIDECASGVTC-----PYNRVCV----- 192
 DB 1216 SDGSYECSC-QPFAIMPORSCCTDIDECEDPNICDGGCCTNIPGEYRCLCYDGFMAE 1274
 QY 193 -----NTEGSYCKCHIGFELYISGRYDCIDINECTMDSH 228
 DB 1275 DMKTCDVNECDLNPNICLSGTCENPKGSFICHCDMGYSK--KGTGCTDINECHIGAHN 1332
 QY 229 TCSHANCENFTGSEFKCKCKGKGYNGGLRCSAIPENS 265
 DB 1333 NCGRHAVCTNTAGSFKSCSSPGMIGDGIKCTDLDECS 1369

RESULT 13
 O88840 PRELIMINARY; PRT; 3857 AA.
 AC O88840:
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MUTANT FIBRILIN-1.
 GN FBN1.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN
 RP SEQUENCE FROM N.A.

RC STRAIN-B10.D2; PubMed=9405934;
 RA MEDLINE=98069008; Bora C.A., Murai C., Casares S., Kasturi K., Nishimura H., Honjo T.,
 RA Matsuda F.;
 RT "Structure of the mutant fibrillin-1 gene in the tight skin (TSK)
 mouse.";
 RL DNA Res. 4:267-271(1997).
 DR EMBL; AF007248; AAC62317.1; -.
 DR HSSP; P35555; IAPJ.
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR002557; Chitin binding.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR002212; TB.
 DR InterPro; IPR000822; Znf-C2H2.
 DR Pfam; PF00008; EGF; 64.
 DR Pfam; PF00683; TB; 12.
 DR SMART; SM00494; Chpbd2; 2.
 DR SMART; SM00179; EGF_CA; 60.
 DR SMART; SM00001; EGF_like; 4.
 DR PROSITE; PS00010; ASX_HYDROXYL; 61.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
 DR PROSITE; PS01186; EGF_2; 50.
 DR PROSITE; PS01187; EGF_CA; 61.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
 KW SEQUENCE 3857 AA; 418301 MW; 5BC0618BC527E04C CRC64;

Query Match 12.5%; Score 383.5; DB 11; Length 3857;
 Best Local Similarity 32.2%; Pred. No. 7.9e-22;

Matches 86; Conservative 34; Mismatches 82; Indels 65; Gaps 11;

QY 39 GVCHYGLIACCYGWRNRSGVCEATCEPG-----CKRGEV- 75
 Db 1129 GICH-----NTEGSYRCECPRGHQLSPNISACIDINCELSANICPHGRVCN 1175
 QY 76 --GPKKCRFCFY---TGKTSQDVNCEGMRPCQHRVNTNGSYKCCLSGHMIMPRA 130
 Db 1176 LIGKQACACNPGYHPTHRLFCVDIDECSINMGCECTFNDSGYSQCPGFALMPD- 1234
 QY 131 TCVNSRTAMINCOYSCEDTE---EGPO-----CICPSSGLRLAPNGRDCLDIDEC 178
 Db 1235 ----QRCTDID---ECEDNPICDGGCCNTNIPGEYKLC-IDGFMASDMMTCVDVNEC 1286
 QY 179 ASGVICYENRCVNTFSGSYCKHIGFELQYISGRYDCIDINECTMDSHTSNHANCEN 238
 Db 1287 DLNPIC-LSGTCENTKSEFICDMGYSK--KGTGCTDINECEIGHANCGRAVCTN 1343
 QY 239 TQGSFKCKCKQGYKNGRLGRCNAIPENS 265
 Db 1344 TAGSFKCSCSPGWDGAIKCTDIDEC 1370

RESULT 14

Q99K58 PRELIMINARY; PRT; 1174 AA.

AC Q99K58; 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE SIMILAR TO FIBULIN 2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MAMARY TUMOR. MAP-TGF ALPHA MODEL. 7 MONTHS OLD, GROSS
 RC TISSUE.;
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; BC005443; AAH05443.1; -.
 DR HSSP; P00736; IAPQ.

DR InterPro; IPR000020; Anaphylatoxin.
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF_Ca.
 DR Pfam; PF001821; ANATO; 2.
 DR Pfam; PF00008; EGF; 6.
 DR SMART; SM00104; ANATO; 3.
 DR SMART; SM00181; EGF; 11.
 DR SMART; SM00179; EGF_CA; 9.
 DR PROSITE; PS01177; ANAPHYLATOXIN_1; 3.
 DR PROSITE; PS01178; ANAPHYLATOXIN_2; 3.
 DR PROSITE; PS00010; ASX_HYDROXYL; 5.
 DR PROSITE; PS01186; EGF_2; 5.
 DR PROSITE; PS01187; EGF_CA; 9.
 DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
 KW SEQUENCE 1174 AA; 126460 MW; 8D628AC710FBAB8 CRC64;

Query Match 12.5%; Score 382; DB 11; Length 1174;
 Best Local Similarity 28.5%; Pred. No. 2.3e-22;

Matches 114; Conservative 39; Mismatches 139; Indels 108; Gaps 20;

QY 47 LACCYGWRNRSG---VCEATCEPG---CKPGE-CV---GPKKCRFCFYGT----GKTC 92
 Db 831 LVCGKRYHANEGSGECVDYNECEFTGVHRCGEGQLCYNLPGSTRCDCKPFGDARGETCI 890
 QY 93 ODVNECGMRP-PCQHRVNTNGSYKCCLSGHMIMPAT-CVNSRTAMINCOYSCEDT 150
 Db 891 -DVNECWSFGRICQHTCENTPGSYRSCAAGFLAADGKHCEDEVNECETRSCQECANI 949
 QY 151 EGGPQCLPSSGLRLAPNGRDCLDIDECASGR-VICPYRRKRVNTFSGSYCKHIGFELQ 209
 Db 950 YSYCYCYC-RQGYQLADEHTCTDIDECAGAGILCTF--KCVNPGSYQACPEQGYTM 1006
 QY 210 YISGRYDCIDINECTMDSHTSCHHANCENFQSGFC---KCKQGY-KGNGLRCSAIPENS 265
 Db 1007 MANGR-SCKDLDICALGTHNCSEAEKCHIQSFRCLRDPCPNYRVSETCERTCO 1065
 QY 266 VKEVLRAPTIDRIKLLAHKNSMKKAKIKNVTPERTPTPKVNILOPNEYELVSRG 325
 Db 1066 ITECOTSPARI---THYQLNFQGLVPAHIFRIGAP----- 1100
 QY 326 GNSHGKKNKEKMEKGELEDEREALKNDIEENSLRGDF---FPKYNEAGEF----- 377
 Db 1101 -----AFADGTLSLTITTKNGEGYFVTRRL 1125
 QY 378 ---GLIVQRKALYSKLEHKDINISVDCSFNHGICDMKQ 413
 Db 1126 NAVTGVSLGR---SVLEPRDFALDVEKML-----WRQ 1155

RESULT 15

Q9BI05 PRELIMINARY; PRT; 2189 AA.

AC Q9BI05; 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MICRONEME PROTEIN 4.
 GN MIC4.
 OS Eimeria tenella.
 OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;
 OC Eimeria.
 NCBI_TaxID=5802;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HOUGHTON;
 RA Tonley F.M., Billington K.J., Bumstead J.M., Clark J.D., Monaghan P.;
 RT "Ethica: A microneme protein from Eimeria tenella that contains tandem
 RT arrays of epidermal growth factor-like and thrombospondin type-1
 RT repeats.";
 RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AJ306453; CAC34726.1; -.
 DR HSSP; P35555; IEMN.

DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR000884; TSPI.
DR Pfam: PF00008; EGF_24.
DR SMART: SM00181; EGF_30.
DR SMART: SM00179; EGF_CA_30.
DR PROSITE: PS00010; ASX_HYDROXYL; 22.
DR PROSITE: PS01186; EGF_2; 18.
DR PROSITE: PS01187; EGF_CA; 18.
DR PROSITE: PS50092; TSPI; 1.
KW Calcium-binding; EGF-like domain; glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 2189 AA; 224388 MW; C272A420B94FCB2D CRC64;

Query Match 12.2%; Score 374.5; DB 5; Length 2189;
Best Local Similarity 32.2%; Pred. NO. 2.1e-21;
Matches 93; Conservative 32; Mismatches 77; Indels 87; Gaps 14;
QY 44 GTKLAC-CY-GMRRNSKGVCEATCEPGCK-----FGE---CV---GPNKCRC 82
DB 966 GTEAICTCHSGYEGNGEG-----EEGCKNIDECYVGEPCKDFGEGYCVDSFGSFGSC 1019
QY 83 PPGYTGKTCG-QDVNEC-GMKPRPCQ--HRCVNTGSGYKCFCLSG----- 123
DB 1020 ATGFIKRRCTCODIDECIDGKMTCAPIYGICTNTVGSFTCSAAGFTGDIETCEDIDEC 1079
QY 124 ----HMLMPDATCYNSRTCAMINQ-----YCEDTEEGPQCLCPG----- 160
DB 1080 ATAAHTCDPNATYCVNTVGSFEGCKEGFSGDGHCTCTDIDE--CADPMLNKCDTHKGIQ 1136
QY 161 -----SGLRLAPNGRDCIDIDECASGKVIQFYNRRCVNTFGSYCKCHIGFELQ 209
DB 1137 NGTSGYTCGCGRPGYSLADGFTCDNVDECAAGTATCGERSFCVDITQSYKCECKNG----- 1192
QY 210 YISGRYCIDINECTMDSHTCGSHANCFNTQGSFCKCKQGYKGNGLRC 258
DB 1193 YRQCGEDCVADVDECAVDVHTCSEHATCTNTGSHTCCTCNEGYQGDGKRC 1241

Search completed: October 17, 2002, 16:01:27
Job time : 40 secs

THIS PAGE BLANK (USPTO)